

The University of Chicago

STATISTICAL DEFLATION IN THE
ANALYSIS OF ECONOMIC SERIES

A DISSERTATION SUBMITTED TO THE FACULTY
OF THE SCHOOL OF BUSINESS IN CANDIDACY
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

1941

By
JOHN H. SMITH

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

1941

The University of Chicago

STATISTICAL DEFLATION IN THE
ANALYSIS OF ECONOMIC SERIES

A DISSERTATION SUBMITTED TO THE FACULTY
OF THE SCHOOL OF BUSINESS IN CANDIDACY
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

1941

By
JOHN H. SMITH

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

1941

Copyright 1941
By JOHN H. SMITH

PRINTED IN THE U.S.A.

PREFACE

Many devices can be used as aids in the determination of net relationships among economic variables. Mechanical adjustments designed to remove the effects of other variables are commonly known as statistical deflation. In this study, other techniques which accomplish the same purpose are also considered in connection with the general problem of statistical deflation. In addition to the familiar division by index numbers these include regression analysis and adjustment for seasonal variation and secular trend.

Throughout the study each deflation adjustment is considered as a step in a regression analysis in which the final goal is the determination of net relationships. This approach avoids discussion of methods for accomplishing one step in analysis without reference to the remainder of the study. Emphasis on the importance of specification facilitates the formulation of criteria for choice among available methods of statistical deflation.

It is a pleasure to acknowledge my indebtedness to Professor Theodore O. Yntema for aid in formulating the problem and for constructive criticisms after reading the manuscript. Professors Neil H. Jacoby and Paul H. Douglas also contributed valuable suggestions. Miss Grace T. Gunn aided me by explaining the studies of marginal productivity which I used as examples. Last, but not least, my wife helped me to clarify my exposition and to eliminate many misleading statements. For any errors or ambiguities which remain, I alone am responsible.



TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
Statement of the Problem	1
The Scientific Method in Economic Research	3
Limitation of the Field of Investigation for this Study	6
The Problem of Specification	10
Probability and Sampling Theory	13
II. THE MEASUREMENT OF TREND	17
The Components of Time Series	17
Trend Adjustment as a Deflation Technique	18
Criteria for Choice among Methods of Trend Measurement	18
Moving Average Trends	20
Graphic Determination of Trend	21
The Method of Least Squares	22
Choice of a Trend Function	26
The Modified Exponential Trend	27
The Method of Selected Points	31
The Summation Method	31
The Differential Equation Method	32
The Method of Moments	32
The Method of Least Squares by Iteration	33
The Variate Difference Method	34
III. ADJUSTMENT FOR SEASONAL VARIATION	36
The Seasonal Component	36
The Method of Monthly Means	38
The Method of Trend Ratios	41
The Ratio-to-Moving-Average Method	43
The Link Relative Method	44
The Logarithmic Second Difference Method	45
Adjustment for Changing Seasonal Variation	47
Graphic Methods of Adjustment for Seasonal Variation	49
Miscellaneous Problems in Seasonal Adjustment	50
Advantages of Seasonal Adjustment	53
IV. ECONOMIC RELATIONSHIPS AND REGRESSION ANALYSIS	56
Abstraction from Important Problems	56
Definition of Net Relationships in Regression Analysis	56
The Mathematical Theory of Exchange	59
Small Unexpected Changes in Quantities	61
Variations in Both Quantities and Prices	65

Chapter	Page
IV. (Continued)	
The Weighted Regression	67
Weighted Regressions for a Simple Specification	69
The Method of Maximum Likelihood and Single-Valued Relationships	71
Numerical Approximations of Parameters of Single-Valued Relationships	76
Reliability of Regression Coefficients	77
Confluence Analysis	80
A Study of Synthetic Production Functions	84
The Measurement of Marginal Productivity	86
V. REGRESSION ANALYSIS OF ECONOMIC TIME SERIES	89
Choice of Components for Use in Regression Analysis	89
Statistical Deflation in Terms of Regression Analysis	96
VI. DEFLATION AND THE MEASUREMENT OF ECONOMIC VARIABLES.	100
National Income as an Example	100
Definition of National Income	100
Net Changes in National Wealth	101
Non-Monetary Income	103
Business versus Personal Expenses	104
The Use of Deflation in the Measurement of National Income	104
Obsolescence	107
Omission of New and Obsolete Goods and Services	108
The Components of National Income	110
Comparisons of Incomes of Classes of Individuals	110
VII. CONCLUSION	114
Statistical Specification and the Scientific Method	114
Specifications for Regression Analysis and Pure Economic Theory	115
Adjustment for Seasonal Variation	116
Trend Analysis as a Deflation Technique	116
Reliability of Regression Coefficients	117
Definition of Economic Variables	118
BIBLIOGRAPHY	121

CHAPTER I

INTRODUCTION

Statement of the Problem

Statistical deflation is the process of adjusting a series for the effects of one or more variables which affect it. For example, the division of a price series by an index number of prices is known as deflation. The purpose of this technique is to remove from the price series the effects of changes in the purchasing power of money in order that the investigator may study the relation of "real price," the adjusted series, to other economic series. Adjustment for the effects of the variable number of days in the month or of one or more changing characteristics of the population may also be made by a deflation process. Adjusted series are expressed in terms of average per calendar or working day or in terms of average per automobile owner or any other member of the population. This includes the familiar per capita basis of reporting data.

The foregoing examples of deflation techniques are all special cases of a more general problem. The essential feature of deflation in this broader sense is the adjustment of a series for effects of one or more variables which affect it. This preliminary adjustment is usually preparatory to studies involving comparison with other series by means of cross classification, regression analysis, or graphic and tabular comparisons. In this sense, deflation is not confined to the process of dividing one series by another; other methods of deflation are often superior.

For the purpose of illustrating important characteristics of the general problem, adjustment of time series for seasonal variation is an interesting example of statistical deflation. It is relatively easy to show that the most effective method of adjusting a series for the effects of weather, customs, and other variables which cause seasonal variation varies from one problem to another. For present purposes, it is sufficient to state that each of the common methods for adjusting series for seasonal

variation is a mechanical process based on statistical averages of characteristics of the series to be adjusted. When a series is to be adjusted for seasonal variation, the method chosen should be the most efficient among methods available for the purpose. It is possible to combine components of several of the methods in common use if special characteristics of the problem make such modification desirable.

When a considerable portion of the seasonal component of a time series is associated with a measureable characteristic of the weather, such as temperature, adjustment by means of a regression function is often more effective than a mechanical adjustment for seasonal variation. For example, the amount of steam required to heat a given building depends upon temperature and other weather characteristics, such as wind velocity, which are highly correlated with temperature. It is possible to adjust a monthly steam consumption series for effects of changing weather conditions by means of the regression function of monthly steam consumption on a measure of temperature deficiency. The most readily available measure of this type is called degree days. For every day the mean temperature in degrees Fahrenheit is subtracted from 65. The sum of the positive differences, for a week or a month, is the number of degree days for the period. Residuals of monthly steam consumption on degree days are used at the University of Chicago for purposes of administrative control of steam consumption.¹ For purposes of control such residuals are superior to steam consumption series adjusted by means of a mechanical adjustment for seasonal variation.

It is possible to combine ordinary methods of adjustment for seasonal variation with the use of regression functions. For example, if there were seasonal variation among the residuals of steam consumption on degree days, these might be adjusted for seasonal variation. Similarly, seasonally adjusted series may be adjusted for deviations from average values of measureable weather characteristics by means of regression functions. When either method of adjustment is used, it should be regarded as a first approximation to be supplemented by the other if necessary. Hence adjustment for seasonal variation is similar to graphic regression

¹William S. Bard, "The Use of Statistical Techniques in the Administrative Control of Steam Consumption," The Journal of Business of the University of Chicago, X (July, 1937), 251-67.

analysis with respect to the use of successive approximations in the derivation of net relationships.

The Scientific Method in Economic Research

The scientific method consists of a body of techniques which are designed to obtain the most useful description of relationships among the phenomena which constitute the subject matter of the science in which they are to be applied. There are many criteria of usefulness in this connection. Among these are dependability, simplicity, and applicability.

In order to be useful, a description of a relationship must have some degree of dependability. This criterion is sometimes called the criterion of **prediction**, but in economic analysis prediction might suggest prophecy with respect to variables other than those between which the statistically determined relationship exists as well as the form and parameters of the relationship under consideration. The criterion of dependability emphasizes the importance of preference for empirical relationships which are apt to endure as contrasted with those which may be largely due to errors in sampling or measurement or, as is still more common in economic analysis, to effects of variables not included in the analysis.

The criterion of simplicity is often referred to as the law of parsimony or Occam's razor. This criterion states that the simpler of two equally adequate explanations is to be preferred. It is important to note that this does not refer to a single explanation in isolation but rather to the system of scientific principles of which the particular explanation is a part.

The criterion of applicability has gained some popularity in connection with the operational viewpoint of nature. According to this point of view, it is highly desirable to define variables in terms of operations. A strict interpretation of this criterion is not a valid basis for condemnation of deductive reasoning in economic theory where the conclusions are really theorems. It is unquestionably useful to deduce the consequences of the operation of a set of economic laws in connection with a hypothetical model even if the institutional setting considered does not contain important characteristics of the real world. The science of economics is, to a large extent, the derivation of conclusions on the basis of such models with more or less frequent discussions of the

manner in which such conclusions would be modified if the models considered were more realistic. In recent years the field has been extended in two directions; more complicated models have been studied and many attempts have been made to determine economic relationships statistically. The latter approach involves variables which are operationally defined, at least by implication. Statistical studies are often more useful than complicated models. The applicability of models is not always commensurate with their difficulty. The two approaches have been described by Moore as follows:

The mathematical approach to economic theory bifurcated at an early point in its course. One road, travelled chiefly by practical men intent upon the exigent business of the day, has led through a dreary region directly to practical, but unrelated, results; the other road, followed mainly by philosophers with primary interest in causes and relations, has ascended to picturesque heights affording distant views of the ensemble of economic activity, but has stopped short amid the enchanting scene and left the explorers in doubt as to what might be the real destination of so promising a beginning.²

The contrast between practical studies of limited scope and extensive mathematical models is further emphasized by Marshall:

It is obvious that there is no room in economics for long trains of deductive reasoning: no economist, not even Ricardo, attempted them. It may indeed appear at first sight that the contrary is suggested by the frequent use of mathematical formulae in economic studies. But on investigation it will be found that this suggestion is illusory, except perhaps when a pure mathematician uses economic hypotheses for the purpose of mathematical diversions; for then his concern is to show the potentialities of mathematical methods on the supposition that material appropriate to their use had been supplied by economic study. He takes no technical responsibility for the material, and is often unaware how inadequate the material is to bear the strains of his powerful machinery.³

Marshall's warning suggests that it is frequently impossible to classify economic phenomena in such a way that measurements based on the individual classes are reliable measures of corresponding measurements in an econometric model even when these measurements are adequate for use in problems less extensive in

²Henry Ludwell Moore, Synthetic Economics (New York: The Macmillan Company, 1929), p. 1.

³Alfred Marshall, Principles of Economics (8th ed.; London: Macmillan and Co., 1930), p. 781.

scope. J. M. Clark says that "one source of confusion lies in the fact that the economist, being a scientist with accurate observation as his dominant purpose and omniscience as his will-o'-the-wisp, tends to give terms the meaning they would have to an omniscient observer, forgetting that if such beings exist they have no need of his analysis."⁴ A particular problem in economic research requires that the classification of economic phenomena be adapted as closely to the purposes of the problem as is feasible if the optimum results are to be secured. Moore's discussion of the problem is interesting in this connection. He says:

A scientific classification, like a scientific method, is to be judged by its fitness for its purpose. In the historic treatment of capital and income much ability has been directed toward a definition of terms and an effort to bring them into a logical classification. But the goal of the great talents thus employed has not always been well defined, and without a distinct specification of the goal such dialectical endeavors might obviously be endlessly prolonged. The Ecole de Lausanne has been exemplarily clear as to its aim. Its ultimate object has been to describe the ensemble of the interrelationships in exchange, production, and capitalization, and with this object always in view it has defined and classified capitals and services.⁵

The foregoing description of the aim of the Lausanne School does not indicate explicitly the form of their description of relationships, but Moore goes on to say:

This problem of description is approached by successive approximations. In the theory of exchange those interrelations are considered which determine the prices of consumable material goods and directly consumable services. In the theory of production those supplementary interrelationships which are required to determine the prices of raw materials and productive services are added; in the theory of capitalization the remaining conditions which are necessary to determine the prices of capital goods are introduced into the general system of equations.⁶

The general system of equations to which Moore refers are the equations which describe the equilibrium of the economic system in terms of economic relationships for which both the functional form and parameters are unknown. It is necessary to impose certain conditions on some of these functions in order to avoid the possibility of multiple solutions and to insure that the re-

⁴ John Maurice Clark, Studies in the Economics of Overhead Costs (Chicago: The University of Chicago Press, 1923), p. 44.

⁵ Moore, op. cit., p. 17.

⁶ Ibid., p. 19.

sulting equilibrium is stable. In other words, the Walrasian equations serve a useful purpose in showing that mutual interdependence among economic variables does not involve circular reasoning and that the number of variables is equal to the number of unknowns. However, classification of economic phenomena of the real world in such a way as to enable an economist to determine statistically the form and parameters of any economic relationship of the Walrasian model is beset with great difficulties.

Any statement of a general relationship or principle implies a classification of phenomena. In many fields of knowledge, especially in the social sciences, the phenomena which are classified are not perfectly homogeneous. In order to derive a body of related principles which make up a particular science, it is necessary to compromise between difficulties associated with an unlimited number of classes and those associated with an increasing degree of heterogeneity. It is always possible to suggest that a heterogeneous class of phenomena should be separated into two or more classes which are more nearly homogeneous. In this connection Preinreich has said:

As Walras has pointed out long ago, the solution of any economic problem ultimately resolves itself into the task of solving all other problems at the same time. In such circumstances, the last-comer is always in a strategic position. No matter how many simultaneous equations his predecessors may have set up, he can always prove their naivete by pointing to another equation, which they ought not have overlooked.⁷

Limitation of the Field of Investigation for this Study

The general problem of deflation in the analysis of economic series is so broad that it is not feasible to consider all aspects of both methodology and subject matter in a single study. Fortunately, portions of the field have been topics for recent critical surveys and many important statistical techniques are too familiar to require extensive explanation. For example, it is hardly necessary to explain least squares regression analysis, and the techniques used in graphic regression analysis are also familiar to students of regression analysis. These and other well

⁷Gabriel A. D. Preinreich, "Annual Survey of Economic Theory: The Theory of Depreciation," Econometrica, VI (July, 1938), 239.

known topics can be summarized briefly with special emphasis on topics with respect to which the writer has some constructive suggestions to offer.

The investigation of the subject matter of any single important field in which deflation is used is a study which rapidly becomes specialized. It seems to be impossible to state a statistical specification of sufficient generality and applicability to form the basis of a useful mathematical investigation of optimum methods of deflation. There seems to be ample justification, however, for a survey of available methods of adjusting economic series for the effects of variables other than those between which the net relationship is to be determined. In such a survey emphasis should be placed on problems which are especially important in connection with economic research.

Detailed consideration of the difficulties of the statistical determination of economic relationships must be postponed but some of the most important ones can be summarized briefly. The equations of equilibrium refer to an economy in which important dynamic changes do not occur. In the absence of such changes the variation necessary for the determination of the form and parameters of a particular relationship is impossible. Because Moore and his followers have been interested in dynamic relationships, they have had to formulate hypotheses regarding the nature of economic phenomena which differ from the stationary state for which the Walrasian equations are adapted. In addition to the necessity of modification of the theoretical model, it is necessary to use index numbers instead of individual series because only a limited number of variables can be subjected to statistical analysis and also because it is practically impossible to determine statistically the separate effects of closely correlated variables.

Combining individual series into composite variables introduces difficulties which are superimposed on those inherent in the measurement of economic phenomena and uncertainty as to the ideal series which the economist would like to have if they were available. According to Clark, "One's approximate estimate is likely to be a great deal better if it is least aimed at the thing one really needs to know than if it is aimed at something else."⁸

⁸Clark, op. cit., pp. 176-77.

It is well known that empirical relationships among time series are apt to be unreliable. Statistical techniques which are valuable aids to scientific inference in fields where the data are not subject to trends, cycles, and seasonal variation often prove to be misleading when applied to time series analysis. Nonsense correlations also arise in cross section analysis of economic series for the same fundamental reason; observations of the various series refer to non-comparable institutional settings. Probably the most common difficulty in such studies is the presence of a size component in two or more of the series. For example, the study of production functions by means of industry totals is hampered by the fact that the average size of establishment varies from industry to industry. The capital, labor, and value added series are also affected by the number of establishments in each industry but the effects of this disturbing element can be removed by a simple deflation technique—dividing by the number of establishments. Unfortunately, it is not possible to eliminate the effects of average size of establishment so easily.

It will later be shown that the introduction of a sufficiently important element of size in a set of observations on value added and factors of production for firms otherwise comparable leads to homogeneous production functions. Thus the evidence value of empirically derived exponents whose sum is approximately unity is greatly lessened. When the effects of heterogeneity are important the same sum would be expected on the basis of any alternative hypothesis.

It is possible to arrive at similar conclusions regarding other disputed points in the analysis of economic series with comparable simplicity, that is, many important principles can be demonstrated on the basis of simple statistical specifications for the case in which the data are complete and accurate. Therefore, this method of approach is followed as much as possible in order to emphasize important principles based on characteristics of economic problems which are not highly mathematical in character. For example, normality of arrays is usually specified in order to make extensive use of the principle that the method of least squares leads to maximum likelihood estimates. As a matter of fact, deviations which must be averaged out in the analysis of economic series are probably more often leptokurtic. Even when arrays are leptokurtic, the method of least squares is often the

most useful method of obtaining first approximations. Allowances for undue weighting of extreme deviations inherent in the method of least squares can be made most effectively in connection with a study of the residuals.

In view of the many types of difficulties which are characteristic of statistical research in economics, it is apparent that careful formulation of problems is necessary in order to avoid meaningless results. In addition to the difficulties inherent in the measurements which form the basis for statistical investigations, there are important difficulties of exposition. One of these is the fact that so many words used in economics have many meanings and shades of meaning both in popular use and in the technical language of economists. Another is connected with the impossibility of repeating all of the qualifications in every sentence. For example, Marshall says, "The question whether Ricardo is to be supposed to have meant nothing by the first chapter of his book, merely because he did not constantly repeat the interpretation clauses contained in it, is one which each reader must decide for himself according to his temperament: it does not lend itself to be solved by argument."⁹

There is no uniformity of opinion among economists as to the place of statistical methods in economic research. Some economists are probably too naive and others too skeptical with respect to the results of statistical studies. Moore's discussion of this subject represents a desirable attitude.

The art of using mathematical methods in the service of science consists in seizing upon those parts of the subject matter that are both critical in importance and quantitative in character, and then discovering relations between these immediately measurable, and other less obviously quantitative, facts. The test of the sufficiency of the mathematical liaison in each individual case is its rendering possible the forecasting of the unknown from the relation with the known. An index of the state of perfection of the science is the degree in which these mathematical connections are extended over the range of the subject matter of the science.¹⁰

The term "forecasting" as used by Moore was probably not intended to suggest that a study whose scope is relatively limited can serve as the sole basis for economic forecasting in the popular

⁹Marshall, op. cit., p. 817.

¹⁰Moore, op. cit., pp. 176-77.

sense, but rather that a change in a series used as an independent variable can be used to "tell us how . . . [the dependent variable] will differ from what it would have been if that cause had not been introduced: and this and not the other, is the information of which we are in search."¹¹

Since empirical relationships among economic series are characterized by deviations from single-valued relationships, tests of significance, confidence intervals, and other devices based on the theory of sampling are sometimes used to support the inference that the statistical relationship represents a causal relationship and to determine limits within which parameters of the underlying relationship probably lie. These devices depend on hypotheses or statistical specifications which describe not only the functional form which is approximately satisfied by the observed series, but also the probability distribution of the deviations from the specified relationship. Such specifications are more or less unrealistic because deviations specified to be random are frequently caused by neglected variables and successive observations (in case of time series) are not independent. Even in cross section studies, observations are rarely homogeneous with respect to the purpose of the study. Specifications simplify discussion, and methods optimum for a particular specification are usually superior to methods which are optimum for other more unrealistic specifications.

The Problem of Specification

Statistical determination of relationships among economic series is usually based on an analysis of several available series, each of which is influenced to a considerable extent by a large number of variables. The list which follows describes some of the characteristics of economic series and limitations of statistical methods which give rise to difficulties in this type of analysis.

1. Most economic variables are related and only a few can be included in one study. Therefore, it is impossible to study a closed system of variables in a single investigation.
2. Efforts to reduce the number of variables by the use of broader classifications result in an increasing degree of heterogeneity.

¹¹A. C. Pigou, The Economics of Welfare (London: Macmillan and Co., 1920), p. 12.

3. Available series are rarely good approximations to variables used in economic models. Even when the series are intended to measure the desired variables, the degree of accuracy of measurement is often low.

4. Statistical determination of certain plausible types of relationship is difficult or impossible. Examples are:

- a) distributed lags with changing periods, and
- b) effects of cumulative changes which give rise to the self-generating business cycle.

5. The functional form of the most useful description of the relationship among a set of economic series is usually unknown and its unknown parameters are usually dependent on the type of institutional setting in which the series are observed. In many cases the optimum form of the relationship varies with changes in the setting over the minimum period for which data for its determination would be sufficient.

6. Slight dispersion or amplitude of fluctuation in related variables during the period studied may affect the reliability of statistical relationships intended for use in future periods. Increase in amplitude or deviation from the average attained during the base period affects adversely the usefulness of a careful statistical analysis.

Emphasis on the importance of avoiding difficulties in connection with each of the foregoing conditions varies from one problem to another. The efficiency of alternative methods with respect to avoidance of each type of difficulty also depends on problem conditions. Hence, the choice among methods should be based on a careful consideration of the following points:

1. The probability of the existence of each condition which might give rise to difficulties
2. The degree of emphasis on the avoidance of each type of difficulty as determined by the purpose of the investigation
3. The efficiency of each alternative method under the conditions of the problem with respect to avoidance of each type of difficulty, especially those which are considered to be most serious

The foregoing treatment of the problem of making a choice among alternative statistical methods for removing the effects of disturbing variables from economic series is sufficient to indicate the nature of the difficulties which are apt to be encountered in making such a choice. Some organizing device is necessary, however, in order to make the discussion more specific. Such a device is suggested by the fact, emphasized in the statement of the general problem of deflation, that the purpose of deflation is to facilitate comparison of the adjusted series with related series.

Comparisons of adjusted series need not take the form of regression studies, but it is convenient to discuss the problem in terms of procedures connected with regression analysis.

Without minimizing the importance of the difficulties listed at the beginning of this section, it is possible to treat the relationship to be determined as a partial regression. The specification of a universe from which available data may be considered to be a sample is also a useful device. When the universe and conditions of sampling so specified are sufficiently simple, it is possible to derive optimum methods for estimating the values of the unknown parameters. Estimates derived by methods appropriate for simple universes and conditions of sampling, like conclusions derived from the theoretical models of the economist, may be used as evidence in the discussion of real problems. It is necessary to consider the nature and extent of revision which would be necessary if the model were more realistic. The usefulness of optimum methods of estimating parameters is limited because manageable specifications rarely correspond closely to problem conditions in the analysis of economic series. In spite of this fact, simple hypothetical universes and conditions of sampling facilitate the discussion of problem conditions and definition of purpose.

Although specification of a universe and of conditions of sampling is primarily a device for facilitating the organization of a problem, its importance should not be minimized. When there are serious discrepancies between hypothetical universes and the conditions of the problem, it is especially important to consider the effects of such discrepancies in connection with the interpretation of the results of an investigation. Koopmans sets forth the following rules regarding the specification of the parent distribution:

1. In the specification should be implied all a priori knowledge on the nature of the data which is relevant to the estimation of the required parameters.
2. In specifying the parent distribution such a posteriori information as to its form should be worked up as may be extracted from the sample itself with a reasonable degree of reliability and is relevant to the estimation of the required parameters.
3. Extension of the specification by the introduction of additional assumptions not imposed by the rules I and II should be avoided as much as possible.

4. The specified form of the parent distribution should be neither so general nor so complicated as to make mathematical treatment of the problems of estimation and distribution too cumbersome and intricate or even impossible.

5. In so far as the specification extends beyond the elements supplied by the rules I, II and III, the accuracy of estimation of the required parameters should not be very sensitive to an inexact fulfilment of the additional assumptions.¹²

Probability and Sampling Theory

The probability that a given statement or proposition is true is the relative frequency of truths in the class of propositions in which the one under consideration belongs. For this purpose, propositions should be classified on the basis of every relevant type of evidence tending to support or refute them. This definition is not operational so far as practical application is concerned because such classification could hardly be applied to any important type of inference. It does serve to indicate some of the important considerations on which the validity of scientific inference depends. It should be noted in the case of propositions defining a range of values in which a given parameter of a universe lies, types of evidence other than sample values of Student's *t* may be relevant as bases of classification. Fiducial probabilities obtained from probability integrals of sampling distributions may be misleading when non-statistical evidence is important. In problems involving economic series it is especially important to consider the plausibility of the specified universe and conditions of sampling and to make adequate allowance for discrepancies between the specification and problem conditions in the interpretation of such probability statements.

The result of any test of significance is a measure of unusualness for a sample statistic from a hypothetical universe under the conditions of sampling specified in the hypothesis. If the statistic is sufficiently unusual as shown by a small probability integral, the hypothesis should be rejected in favor of a plausible alternative hypothesis under which the measure of unusualness would not be so small. The most effective way to present a test of significance for a statistic in an attempt to establish the truth of a hypothesis is to show that the probability

¹²T. Koopmans, Linear Regression Analysis of Economic Time Series (Haarlem: De Erven F. Bohn N. V., 1937), pp. 8-9.

integral of the statistic on the basis of the most plausible alternative hypothesis is small. For example, if a test of significance is to be used to support the hypothesis that the regression coefficient of y on x is positive, the probability integral of a positive sample regression coefficient should be obtained on the basis of a hypothesis which states that the universe value is zero. Such a hypothesis is often called a null hypothesis. Neglecting the possibility that other parts of the hypothesis should be changed, a sufficiently small measure of unusualness obtained from such a test should be considered as good evidence that the regression coefficient in the universe is positive. If the null hypothesis is rejected on the basis of the test, it must be in favor of the hypothesis which is to be established since the measure of unusualness would be still smaller if the universe value of the regression coefficient were negative.

It is important to note that the value of a probability integral as evidence in support of an inference is much less if the null hypothesis which was set up merely to be disproved is not the most plausible alternative hypothesis. A good example of such a hypothesis is one which is tested in harmonic analysis of economic time series. The hypothesis on which available tests are based is that the items of the series constitute a random sample from a normally distributed universe. The high positive serial correlation in most time series, especially those which are apt to be harmonically analyzed, eliminates this hypothesis from serious consideration as a plausible one.

Although tests of significance in harmonic analysis are not of great interest in themselves, harmonic analysis resembles other types of analysis of economic series in that sampling distributions based on plausible hypotheses are not available. Hence, in spite of increasing interest in specification in recent years, tests based on sampling theory are of limited usefulness in connection with the study of economic time series. It is especially important, therefore, that research workers in economics should be familiar with the shortcomings of the available methods.

Many sample statistics are members of more than one sampling distribution and it is important to consider each statistic as a member of that sampling distribution which is most appropriate for the purposes of the problem. For example, the chi-square test of goodness of fit of normal curves fitted to frequency distribu-

tions is often applied using one less than the number of class intervals as the number of degrees of freedom even when the mean and the standard deviation are estimated from the sample.¹³

The objection to the test in this form is that it results in the exclusion of extreme values of test criteria. A statistic derived from a sample whose mean and standard deviation agree with those of the specified universe is compared with the percentile points of a sampling distribution in which there is no restriction on sample means and standard deviations. Less than one out of twenty sample values of chi-square will exceed the five per cent point because values of chi-square obtained from samples in which the means and/or standard deviations differ from the corresponding values in the universe are larger on the average than those obtained from samples in which the means and standard deviations agree with those of the universe. The misuse of the chi-square test is an excellent example of the common practice of comparing restricted test criteria with percentile points of sampling distributions from which extremely large or small values are not excluded.

The number of degrees of freedom for the chi-square test of goodness of fit can be adjusted by subtracting the number of estimated parameters, but appropriate adjustments for other devices which exclude extreme values of test criteria are not known. Such exclusion usually results from specifying the form of the universe after studying the characteristics of the sample. This is merely an undue extension of Koopmans' third rule of specification which was previously quoted with approval. Exclusion of extreme values of test criteria which results from the application of this rule is often unimportant. For example, the practice of considering relationships between two variables in terms of both logarithms and original measurements should not be discouraged because it is possible to show that more than one in twenty regression coefficients between uncorrelated variables will exceed the five per cent point if the most "significant" of these two types of regression is adopted for each sample. Considering a large number of independent variables and retaining only those which prove to be statistically significant is a more serious misuse of

¹³Frederick C. Mills, Statistical Methods (Rev. ed.; New York: Henry Holt and Co., 1938), pp. 626-27.

the device of "scientific fishing expeditions." The exclusion of small absolute values of test criteria is apparent even when the specification on which the tests are based is moderately realistic. It is even more serious in regression analysis of economic time series.

CHAPTER II

THE MEASUREMENT OF TREND

The Components of Time Series

The specification that a set of observations on a group of related economic time series constitute a random sample from a multivariate universe would be extremely unrealistic. The successive items in most economic time series are positively correlated. This serial correlation is usually accounted for in discussions of time series by the specification that there are four components in each series: trend, seasonal, cyclical, and random or residual components.

In the general method which is most commonly advocated for the study of related time series, trend and seasonal components of each series are measured explicitly by one of several methods to be discussed in detail later. These components are then removed; that is, some function of the series which does not contain these components is constructed. The ratio to trend less the index of seasonal variation is one example of such a function.

When all of the series have been adjusted for trend and seasonal variation, regression analysis is applied to the adjusted series of cyclical residuals. This is appropriate when the relationship to be studied involves only cyclical residuals. Although it is true that many studies in economic research are made for the purpose of improving current explanations of economic phenomena not found in the stationary state, it would be a mistake to infer that all relevant relationships are independent of trend and seasonal components. Even when a research worker is primarily interested in relationships among cyclical residuals, it is necessary to consider the nature of the component of each series which is operationally defined to be a cyclical residual in terms of the statistical techniques applied. Operational definition is not sufficient to identify such a component with the cyclical component of a similar series in an econometric model.

Trend Adjustment as a Deflation Technique

Just as adjustment for seasonal variation is the process of eliminating approximately the effects of weather conditions, peculiarities of the calendar, etc., so the adjustment for trend as a deflation technique is the process of removing the "effects of factors which, though unknown or not admitting of measurement, change but slowly and smoothly with time."¹ As the quotation suggests, trend adjustment is more useful when the effects of neglected variables on the series to be adjusted can be represented by a fairly smooth function of time than it is when such effects are subject to large discontinuities. Examples of neglected variables whose effects are often eliminated statistically by means of trend adjustment are population growth and changes in tastes or in methods of production. The effects of increases in population can sometimes be eliminated satisfactorily by dividing each series to be adjusted by the appropriate population series. This is true when each series to be adjusted is proportional to a certain class of the population and intervals between complete censuses are sufficiently short. Difficulties arise from lack of availability of data and also from changes in tastes and in technology which often accompany changes in population. Hence, trend adjustments are often used to eliminate statistically the effects of population growth even when per capita series might be used.

When available information is not sufficient for determining which type of trend provides the best index number of the effects of variables not included in the analysis, it is necessary to make a choice among alternative methods of trend measurement by means of criteria based on the properties of the trend function with reference to a single series.

Criteria for Choice among Methods of Trend Measurement

The function chosen to represent secular trend should follow the long-term movements of the series without being unduly affected by the short-term movements. It is difficult to clarify this rather vague criterion concerning the appropriate degree of flexibility of trend functions. What constitutes a long-term

¹Henry Schultz, The Theory and Measurement of Demand (Chicago: The University of Chicago Press, 1938), p. 182.

movement is a decision which must be made by the individual investigator. It is not even advisable to advocate trend adjustment as a technique to be used in every problem involving time series. The usefulness of trend adjustment depends on whether studies involving adjusted series are apt to make more significant contributions to existing knowledge. For example, sales of automobiles is an increasing function of population. A regression analysis of unadjusted series of this type would be apt to add little to existing knowledge, especially when changing tastes of consumers and changing quality of product is so important as in this example. On the other hand, trend adjustments may combine with adjustment for population growth approximate adjustments for technological progress and changes in tastes. Hence, residuals from trend might be more useful in analysis than either the unadjusted series or automobile sales expressed on a per capita basis.

Although it is often convenient to make a distinction between the trend of a time series and its cyclical component for purposes of analysis, such distinction is not absolutely necessary. When it is made, it requires the same type of pragmatic justification as does the device of trend itself. For this reason the changes measured by trend need not be permanent; they may even be cyclical in nature. In fact, methods of trend adjustment are sometimes used to eliminate the effects of variables connected with the business cycle. However, if a component associated with the business cycle is specified to be superimposed upon secular trend, the trend function should not be so flexible as to be unduly influenced by cycles.

After a trend of a given series has been measured the series is adjusted for trend by constructing a function which does not contain a trend component. The two most familiar functions of this type are ratios to trend and deviations from trend.

In addition to the criterion of the appropriate degree of flexibility already discussed, it is desirable that methods of trend measurement be simple and objective. Simplicity is desirable not only because it permits saving of time and effort but also because it facilitates, at later stages in the analysis, a consideration of possible effects of trend adjustment on empirical relationships among adjusted series.

Results obtained by using objective methods are apt to be

more readily accepted (by other investigators) because the results are popularly supposed to be less dependent on the subjective opinions of the research worker. This criterion of objectivity is difficult to apply because there are subjective decisions to make in the application of any method. One investigator might make a given decision and consider it objective while another investigator who is more familiar with the method might make exactly the same decision and consider it subjective. Cases in which "objectivity" is a function of experience are discussed in connection with some of the particular methods of trend measurement.

Moving Average Trends

A simple moving average of an odd number of observations can be used to measure trend. Such a moving average trend has many interesting properties. A moving average is independent of a regular cyclical component whose period is contained as a factor in the period of the moving average. A simple moving average is biased in the direction of concavity when the trend is curvilinear. Large changes in a series cause changes in the moving average at points approximately a half period in each direction. Finally, a moving average is not defined for several observations at the beginning and at the end of the series and, of course, it cannot be used for extrapolation purposes.

In practice, the period of the moving average can seldom be chosen so as to eliminate cyclical fluctuations perfectly because business cycles are irregular in both amplitude and period. The choice of period for a moving average is often based on comparative advantages of long and short periods. A moving average of sufficiently long period eliminates a large part of the cyclical component whether it is regular in amplitude and period or not. On the other hand, when the period is short, the trend is defined for a larger number of observations and the bias in the direction of concavity is less serious in case the trend is curvilinear.

When the curvature of the trend is small the bias of a simple moving average is unimportant. When the curvature is great, such bias can be partially eliminated by introducing a system of weights.

The device of moving averages is not limited to simple or unweighted averages. Various types of weight patterns can be used to give to the resulting moving average any of several properties

which may be desirable in special problems. Appropriate weight patterns minimize the effects on the moving average of adding or dropping unusual observations and reduce the bias of simple moving averages in the direction of concavity in the case of curvilinear trends or long cycles. The device of weighted moving averages is discussed in detail by Macaulay² and Sasuly.³ Perhaps the most effective way to improve a moving average trend is to smooth it graphically. Graphic smoothing minimizes the discontinuous effects on the trend of abrupt changes in the series. It also serves to define the trend for all observations and may even be used for extrapolation purposes subject, of course, to a loss in objectivity.

Graphic Determination of Trend

Free hand curves used to represent secular trends are perhaps most reliable when they have some objective basis such as a moving average which may have either a simple or complex weight pattern. The device is not limited, however, to graphic smoothing of moving averages. It can be used alone or to supplement any other method of trend measurement. The use of free hand curves permits the investigator to make use of any information he may have with respect to the effects of neglected variables. It is also possible to make direct application of other criteria. Even when supplemented by the use of moving averages and similar devices, graphic smoothing has the disadvantage that its effectiveness depends on the knowledge and skill of the investigator. It is, in a certain sense, more subjective than other methods in common use. For this reason some statisticians prefer methods by which any two competent investigators would obtain the same measure of secular trend for the same series. However, when graphic smoothing is used with full recognition of its disadvantages, it is one of the most practical methods of measuring trend.

If there were a set of objective criteria for choosing the form of trend function, it would be possible to determine trends without making any subjective judgments. However, the im-

²Frederick R. Macaulay, The Smoothing of Time Series (New York: National Bureau of Economic Research, 1931).

³Max Sasuly, Trend Analysis of Statistics: Theory and Technique (Washington, D.C.: The Brookings Institution, 1934).

portance of objectivity has been greatly overemphasized. The differences between free hand trends of a given series as determined by two different investigators are often less important than the differences between trends fitted by the same investigators by mathematical methods. The choice in one case is just as subjective as in the other. In addition, the choice among available mathematical functions must be made in ignorance of the particular component of the series which would be identified as trend by each of the various functions. This is especially true when trends are fitted by inexperienced research workers for whom mathematical methods are most strongly recommended.

The Method of Least Squares

A mathematical trend function which is linear in its parameters can be fitted by the method of least squares in such a way that the sum of squares of the differences between observations of a given time series and the corresponding trend values is a minimum. If the observations of a time series are specified to be a random sample from a universe characterized by a linear trend function and normal homoscedastic arrays of the dependent variable, sample coefficients of a least squares trend function have the maximum likelihood property. This would be a strong argument in favor of the method of least squares for the determination of trend if the specification were realistic, but the high degree of serial correlation which exists in most economic time series is sufficient in itself to make the specification untenable. Hence coefficients of trend functions fitted by least squares should not be considered to be the most probable values of the unknown parameters of the true underlying trend but merely as the coefficients of the particular function of a specified type from which the sum of squares of residuals is least.

Perhaps the most effective way to show how cyclical fluctuations affect the statistical determination of trends is to compare the coefficients of a linear trend fitted by the method of least squares with the corresponding coefficients of a simple synthetic time series. The series contains a linear component on which is superimposed a cyclical component whose booms and depressions are symmetrical about their respective peaks and troughs and are identical except for sign. The origin is chosen at the midpoint of the series so that the intercept is identical with the

mean of the series. It is evident that the addition of any number of complete periods of the type described will not affect the mean of the series. The constant term will differ from the corresponding parameter of the constructed series if, and only if, the number of positive observations in the cyclical component is greater or less than the number of negative observations. Even in the case of irregular cycles the intercept at the midpoint is affected only if the sum of the terms of the cyclical component is not zero.

The slope of a linear trend fitted by least squares is a linear function of the observations where the coefficients are proportional to the distances of the corresponding time points from the midpoint. Thus it is unaffected by the addition of a regular cyclical component which is symmetrical with respect to the origin; for example, one of the type described above which has a peak or trough at the midpoint of the series. If the series begins and ends at normal levels following depressions, the slope will be biased negatively because the negative deviations near the end are multiplied by the largest positive coefficients while the positive deviations near the beginning have the largest negative multipliers. Both of these influences operate to reduce the algebraic value of the slope of the trend. Similarly, the slope is algebraically too large when the series begins and ends at normal levels following booms. More generally, the effect of the cycle on the slope of a linear trend is proportional to the sum of the products of terms of the cyclical component by the distances of the corresponding time points from the midpoint of the series. Thus observations near the beginning and the end of the series are weighted most heavily in determining the slope by the method of least squares. A research worker can frequently determine the direction of the effects of the cycle on the slope by inspection of the signs of the first few and the last few products when these signs are alike even when the number of periods is not integral and the cycles are not perfectly regular. When the cyclical component contains an integral number of smooth regular cycles, the slope is too small if the midpoint of the series is in a descending phase and too large if the midpoint is in an ascending phase of the cycle. When the cycles are neither regular nor integral in number, it is necessary to concentrate attention on product signs at extremes of the series.

The foregoing treatment of the effects of cycles on linear trends is sufficient to show that cycles may have an appreciable effect on such an inflexible function as that represented by a straight line. Effects of cycles on more flexible functions such as polynomials of higher degree are even more important. In fact many statisticians are reluctant to use polynomial trends for this reason. Research workers who use polynomial trends of second degree or higher do so because such trends describe the long run movements of the series more satisfactorily than other functions of comparable simplicity which have been considered. Presumably they would abandon the use of polynomials in favor of any other functional types which would appeal to the investigators or to their readers as being superior for the purpose of description and possibly for extrapolation also. If this appraisal is a fair one, a rational choice among familiar trend functions can be made by studying the available series, deciding the optimum degree of polynomial trend to fit, and comparing such a polynomial trend with the trend obtained by the use of the most satisfactory alternative function. A brief examination of this method will show that there are many arguments for and against some of its component procedures which are worthy of analysis.

It might seem that there could be no objection to the choice of a trend function on the basis of a study of the characteristics of the available series. However, it was shown in the preceding chapter that the choice among methods of analysis on the basis of statistical significance leads to exclusion of extreme values of test criteria. For the case of economic time series this is usually trivial in comparison with other conditions which make unrealistic the specifications on which the common tests of significance are based. For example, it has been shown that the coefficients of the trend of a synthetic series made up of a linear trend and a smooth regular cyclical component can be determined exactly by least squares only if there are an integral number of periods and a peak or a trough occurs at the midpoint of the series. Since close approximations to the conditions cited are rarely encountered in the analysis of economic series, it is necessary to consider the effects of the cycle on trend determination. Similar difficulties which are encountered in connection with seasonal adjustment are usually avoided by fitting trends to annual data.

Statistical methods can be used to determine exactly the parameters of trend, cyclical, and seasonal components of certain simple types of synthetic series. When the constructed series contain random components of the appropriate form, these methods result in estimates which have the maximum likelihood property. Tests of significance, confidence intervals, and other devices based on sampling theory are applicable to the analysis of series which can be described by specifications leading to manageable synthetic series, but sampling tests as they are commonly used are not appropriate even for simple series. Estimates of sampling variances of functions of trend coefficients are usually based on the variance of residuals from trend. They should be based instead on the variance of the random component. These sampling variances refer to sampling distributions of statistics which are not affected by any uncertainty as to the constancy or regularity of the cyclical and/or seasonal components because the forms of these periodic components are part of the statistical specification. Estimates of the parameters of the various components are sample statistics whose sampling variances are functions of the variance of the random component, but in practice, effects of irregularity in the cyclical component are usually more important than sampling errors based on the random component of the series. Hence, formulas for sampling variances which would be appropriate for simple synthetic series yield results which are too small to be consistent with the variance among trend coefficients fitted to sections of typical time series.

In an experiment designed to test the adequacy of the standard error of a linear trend function, Davis found that even the use of the variance of residuals from trend instead of the variance of the random component was not sufficient to account for the variance of errors in forecasts obtained from a series of prices of rail stocks.⁴ This result is not surprising because the cycles in the price series used by Davis are not consistent with the specification on which the formula for the standard error of the function is based. Davis suggested a modification of the

⁴Harold T. Davis, "Significance of Analysis of Variance of Time Series," Abstracts of Papers Presented at the Research Conference on Economics and Statistics Held by the Cowles Commission for Research in Economics (Colorado Springs, Colo.: Colorado College, 1936), pp. 99-102.

formula for the standard error of the function which is approximately consistent with the results of his experiment. He did not include the specification for which it would be appropriate, however, and it seems fair to assume that it is not based on any definite universe and conditions of sampling. No available theory is based on the effects of an irregular cyclical component on trend determination or uncertainty of the form of the function whose parameters are to be estimated. Hence, it is probably better to use the familiar formulas than to use modifications derived without specific reference to the conditions which are responsible for the lack of applicability of the simpler formulas. The familiar formulas apply with precision to an unrealistic specification while arbitrary modifications do not apply to any known universe and conditions of sampling.

Choice of a Trend Function

The function chosen to represent the trend of a series should be capable of following the long run movements. Sometimes the required properties as determined from the appearance of a free-hand trend can be stated in the form of a differential equation. In such cases the solution is used to determine the optimum form of the trend function. For example, the property of a constant relative rate of change is expressed by the differential equation which states that the derivative is proportional to trend. This condition leads to an exponential trend function. Since this function is linear in terms of logarithms of the series, the advisability of carrying out further analysis in terms of logarithms should be considered. Similarly, the use of reciprocals and other transformations should be considered whenever they simplify the analysis.

When a free-hand trend has two minimum points and one maximum point, it would be reasonable to choose a trend function which has the same number of bend points. Since the maximum number of bend points in a polynomial trend is one less than the degree of the function, a polynomial trend could have the required number of bend points only if its degree were at least four. However, the trend function as determined by least squares under these conditions might have less than three bend points or the location of the bend points might differ significantly from those obtained by the graphic method. The most interesting cases occur when the

discrepancy between the mathematical and the free-hand trends appears to be caused by cycles or by restrictions of a mathematical nature. Any investigator who identifies approximately the components of a series which he considers to be trend and cycle is inconsistent when he tolerates important modifications either because of mathematical restrictions on the possible shapes of a given functional type or because the system of weights implicit in the method of least squares permits a strategically located cyclical fluctuation to have undue influence on the determination of a trend function. When confronted with such a dilemma, a research worker should consider the advantages of arbitrary subjective determination of the cyclical component over statistical determination of the same component in partial ignorance of the restrictions imposed by the choice of a particular trend function. These advantages are partially offset by increases in objectivity, greater reader acceptance, and other possible advantages connected with properties of least squares trends. Of course, the use of alternative trend functions should also be considered.

The Modified Exponential Trend

One reason for the use of polynomial trends is the lack of suitable alternative functional types. Another is the fact that the nature of objections to an important class of alternative functions is not clearly understood. A function of this class cannot be reduced to linearity by a simple transformation. Hence, it is not possible to minimize the sum of squares of deviations from trend with respect to all of the parameters in a single solution. However, it is possible to assign arbitrary values to one, or a few, of the parameters in such a way as to permit the determination of the remaining parameters by the method of least squares. An important function of this type is the modified exponential trend which can be written $y = a + bc^x$. This function includes the exponential trend as a special case when the value of zero is assigned to a and it approaches the simple linear trend as a decreases without bound. When a is assigned any arbitrary value less than the minimum value of y , b and c can be determined by least squares by transposing a and taking logarithms. Such a method of trend determination does not minimize the sum of squares of deviations of values of y from the corresponding trend values. For this reason it has been contended that statistics computed in

this way are not true least squares results, that no standard errors can be computed for them, that tests based on sampling theory cannot be applied to them, and that they do not have the maximum likelihood property. None of these objections is of great importance because of the known inadequacies of manageable sampling specifications for series in which the effects of irregular cycles are important, but they are easily answered in terms of sampling theory neglecting the effects of cycles.

When a test of significance is applied to deviations of sample statistics from hypothetical values, probability integrals obtained indicate the relative frequency with which the deviations would be exceeded in samples selected from a given universe under specified conditions of sampling. Coefficients of least squares trends have the maximum likelihood property and all of the common tests of significance are applicable when the observations of the time series are selected at random from normal homoscedastic arrays whose means are described by a mathematical function of the same form as the one used for the trend function. Formulas for the sampling variances of trend coefficients derived under these conditions are unbiased estimates of the variances of normal sampling distributions of the corresponding coefficients or functions when observations in successive samples are taken from the same arrays. Thus coefficients of an exponential trend may have the same properties as those of a linear trend if the foregoing specification is applied to logarithms instead of original observations. Similarly, the specification might be applied to the logarithms of $(y-a)$ where the value of a is chosen arbitrarily in which case the other two coefficients in the modified exponential trend would also have the same properties. There might be some objection to these latter applications if original observations were more nearly homoscedastic than arrays of either of the suggested transformations. Although this objection depends on a condition which varies from one problem to another, there are some reasons for believing that homoscedasticity of relative deviations implied by the logarithmic transformation would be of more general applicability in economic time series. In most cases the distinction would be a trivial one. Another objection is concerned with variation in the coefficient to which arbitrary values are assigned. If the arbitrary value is independent of the sample being analyzed, there can be no argument on this point; it is a parameter

of the specified universe which is not subject to sampling errors.

Consider a large number of series constructed by adding random selections from a normal universe to successive ordinates of a modified geometric trend in which the constant term has a finite negative value. When an arbitrary negative value is assigned to $\underline{a}$ in order to determine values for $\underline{b}$ and $\underline{c}$ by least squares, the intercept referred to the midpoint as the origin is the sum of $\underline{a}$ and $\underline{b}$. Such an intercept would ordinarily be confined to the comparatively small range between the corresponding intercepts of linear and exponential trends fitted to the series. If the values of the observations of the series are large in comparison with deviations from trend, the sampling variance estimated for $\underline{b}$ when an arbitrary negative value is assigned to $\underline{a}$ will approximate closely the sampling variance of the sum of $\underline{a}$ and $\underline{b}$ when all coefficients are determined by least squares by an iterative process. Under the latter conditions the negative correlation between $\underline{a}$ and $\underline{b}$ in their simultaneous sampling distribution is nearly perfect. The sampling variance of their sum, which is much less than the sampling variance of either, is approximately the square of the standard error of the mean of the series. Thus the practice of modifying a geometric trend by the introduction of an arbitrary constant in order to introduce the desired degree of curvature into the trend function also serves under favorable conditions to make the easily computed standard error of $\underline{b}$ approximate the much more complex standard error of the function at the midpoint of the series. This increased applicability of simple standard errors is an advantage, especially when, as is usually the case, the specification on which the sampling tests are based is so unrealistic that computation of a complicated formula is hardly worth the effort.

When a trend function which is linear in its three parameters is fitted by least squares to series in which the common tests of significance are applicable, the sum of squares of deviations from trend is proportional to chi-square where the number of degrees of freedom is three less than the number of observations. When the value of one parameter is known and the trend function is fitted by determining least squares values for the other two coefficients, the sum of squares of deviations from trend is also proportional to chi-square but with one more degree

of freedom. Since this sum of squares is divided by the number of degrees of freedom to derive an unbiased estimate of the variance of the random component, a bias would be introduced if the least squares value of one coefficient were arbitrarily assigned to it in the manner which would be appropriate if it were a known parameter of the universe. In the case of functions which are linear in their parameters the appropriate adjustment is to reduce the number of degrees of freedom by one for every coefficient to which an arbitrary value is assigned when such a value is sufficiently close to the least squares value. Similarly, such an adjustment might be applied in the case of modified geometric trends when an arbitrary value is assigned to one of the coefficients. In view of the many conditions which tend to invalidate arguments based on sampling theory in connection with the determination of trend of a typical economic time series, such a reduction of the number of degrees of freedom should be sufficient.

The method of least squares is probably the most popular method used in the statistical determination of trend functions. For simple functions which are linear in their parameters it furnishes a comparatively simple and objective method of determining the coefficients in such a way that the deviations from trend have several properties which are usually desirable for use in further analysis. The sum of squares of deviations from a least squares trend is less than the corresponding sum of squares based on any other trend function of the same form but with different constants. If the trend function has a constant term, the sum of the deviations is zero. Since most trend functions do contain such a term, deviations from trend are usually expressed as deviations from their mean. When a least squares trend is fitted to a simple transformation of the original observations, the deviations of this transformation of the original observations from its trend have the same properties which deviations of the original observations have when no transformation is used. For this and other reasons further analysis is often carried out in terms of such transformations. For example, in many studies involving groups of related economic time series some or all of the variables are expressed in terms of logarithms.

One disadvantage of the method of least squares is the lack of conformity of the curvature of fitted functions with the degree of curvature which is regarded as desirable under problem

conditions. This is due in part to the properties of polynomials which are so often chosen as trend functions. The modified exponential trend has been discussed as an example of the types of transcendental functions whose properties with respect to curvature in particular regions are more easily predictable than the corresponding properties for polynomial trend functions. The use of a transformation of the dependent variable in the fitting process is often desirable, but it is not always necessary. Logarithmic, trigonometric, and other transcendental functions of time and even such functions as square roots of expressions involving time measured from a fixed point can be used to insure a given degree of curvature or other properties which are desirable in a particular trend function. Functions of this type involving constants with arbitrarily assigned values can be used in linear combination in the same manner as the terms of a polynomial trend.

The Method of Selected Points

When the mathematical function used to describe the trend in a given series is not linear in its parameters, methods other than least squares are often chosen. Of these, the method of selected points is generally simplest. The constants are determined in such a way that the graph of the trend passes through a set of points selected by the research worker. Unless there are definite criteria for the selection of the set of points, the method of selected points is subjective and it has been the subject of criticism for this reason. Sometimes the series is divided into as many sections as there are parameters in the trend function and the trend function is fitted to the section averages centered in the appropriate sections. The process of fitting a linear trend in this way is known as the method of semi-averages. It is interesting to note that the slope of a semi-averages linear trend is less affected by extremely high or low values near the beginning or the end of the series than is the slope of a similar trend fitted by least squares.

The Summation Method

When a series is divided into as many sections as there are coefficients in the trend function, the section sums of the observations or of simple transformations of the observations can sometimes be equated to the corresponding sums of trend values.

This method is often used for fitting the modified exponential, the Gompertz, and the logistic trend functions. For these three important functions which are not linear in their parameters, the sums of observed values, logarithms, and reciprocals, respectively, are equated to the corresponding sums of trend values by sections and the resulting equations are solved for the unknown constants. Of the many available treatments of this method the one by Croxton and Cowden is probably the simplest.⁵

The Differential Equation Method

Hotelling has suggested that coefficients of certain trend functions can be determined with greater accuracy by fitting the function in its differential equation form to the finite differences of the observed time series.⁶ The constant of integration can then be determined in such a way as to make the integrated function pass through any desired point. In particular, the trend function should agree exactly with the most recent observation when the trend is to be used for purposes of extrapolation. A forecast is then based on the rate of change which is estimated more accurately by this method than by others in common use under conditions discussed by Hotelling. These conditions would seem to be more nearly applicable to population growth than to economic time series which are affected to an important degree by the business cycle because the observed first differences are not reliable measures of the long run trend increment for such series. The differential equation method is worthy of consideration for the actual trend-fitting process while the device of basing extrapolations on the most recent observation and the estimated rate of change might well be used even when other methods are used to determine the trend coefficients.

The Method of Moments

There are two general methods of fitting trends to time series by the method of moments. Both methods involve the solution of a set of equations, one equation for each coefficient of

⁵Frederick E. Croxton and Dudley J. Cowden, Applied General Statistics (New York: Prentice-Hall Inc., 1939), chap. xvi.

⁶Harold Hotelling, "Differential Equations Subject to Error, and Population Estimates," Journal of the American Statistical Association, XXII (September, 1937), 283-314.

the trend function. In the first method the i th equation is formed by multiplying each of the observations by the $(i-1)$ th power of x and equating the sum of such products to the corresponding sum of computed trend values in terms of the unknown coefficients. For the case of ordinary polynomial trends these equations are identical with the normal equations used in the method of least squares.

The method involving only finite sums is sometimes modified by substituting for the sums involving the computed trend values the corresponding definite integrals. This process is usually rather simple but difficulties are encountered in the appropriate adjustment of the corresponding sums involving the discrete observed values of the series. The adjustment may consist of a simple multiplication of each sum by the common time interval between successive observations or an interpolation formula may be used to compute the moments of the observed series.⁷ The use of the method of moments for trend determination is usually limited to functions which are difficult to fit by means of familiar methods which have been discussed previously. One such function is the Gompertz-Makeham trend function.⁸

The Method of Least Squares by Iteration

When a trend function is not linear in its parameters, the sum of squares of deviations from trend cannot be minimized with respect to all of the coefficients by the solution of a simple set of normal equations. In many cases, least squares solutions for the coefficients of such trend functions can be closely approximated by an iteration method.⁹ Approximate values of the coefficients are obtained by the method of selected points or some other comparatively simple method. The trend function is then expanded in a Taylor's series about the approximate values in terms of corrections to be applied to them. The sum of squares of deviations from trend is minimized with respect to terms of

⁷Karl Pearson, "On the Systematic Fitting of Curves to Observations and Measurements," Biometrika, I (1901-1902), 265-303.

⁸Henry Schultz, "An Extension of the Method of Moments," Journal of the American Statistical Association, XX (June, 1925), 242-45.

⁹Henry L. Rietz (ed.), Handbook of Mathematical Statistics (Chicago: Houghton Mifflin Company, 1924), p. 68.

the first degree in the corrections. The second approximations to the least squares values of the coefficients of the trend function are then constructed by adding the corrections to the first approximations and the process is repeated to obtain successive approximations. If the first approximations are so close that the terms of second and higher degrees in the corrections are negligible, the process is fairly satisfactory for problems in which least squares solutions are desirable. Since there is rarely any important theoretical argument in favor of least squares trends, and since extra effort is required to approximate least squares solutions when trend functions are not linear in their parameters, simpler methods should ordinarily be used.

The Variate Difference Method

The variate difference method is based on studies of the successive differences of a time series or of two related series. One of its purposes is to estimate variances and covariances of components of the series which are random with respect to time. Successive differences are computed and the ratio of the variance or covariance based on the last differences to the corresponding variance or covariance based on the differences of next lower order is determined. When this ratio is sufficiently close to the limiting value of a similar ratio derived from a long random series the variance or covariance of the last difference divided by an appropriate constant is taken as an estimate of the required variance or covariance. The work of previous writers has been surveyed by Tintner.¹⁰

As presented by Tintner, the variate difference method leads to a trend function in the form of Sheppard's smoothing formula, a weighted moving average which is equivalent to a series of midordinates of polynomial trends fitted to successive short sections of the series. Suggestions are given concerning the degree of the polynomial and the length of sections to which the function is fitted.

Simple statistical specifications, especially those from which the common tests of significance are derived, are usually based on random sampling from a normal universe. Hence, it is

¹⁰Gerhard Tintner, The Variate Difference Method (Bloomington, Indiana: Principia Press, 1940).

not surprising that randomness and normality of deviations from trend are advocated as criteria for choice among methods of measuring and eliminating trend. These criteria would be harmless in themselves if it were not for the fact that they may serve to distract attention from more important criteria. Suppose a given time series is one of a set of series among which relationships are to be studied and one of the common tests of significance is to be applied. The conditions peculiar to time series which lessen the usefulness of the test are not lack of randomness with respect to time of deviations of single series from trend but lack of applicability of other important characteristics of the specification. Among these are (1) the unchanging nature of the universe and (2) the choice, for the various series, of the particular components involved in the multivariate relationship. The criterion of randomness of deviations from trend with respect to time is often in conflict with the principle that it is desirable to study the components among which causally significant relationships can reasonably be assumed to exist.

CHAPTER III

ADJUSTMENT FOR SEASONAL VARIATION

The Seasonal Component

Most economic time series whose observations are less than a year apart are characterized by seasonal variation. Hence a seasonal component is usually specified when such a series is to be analyzed into its components. Since the seasonal variation of most time series is related to temperature, the series of mean monthly temperatures is a good example with which to illustrate some of the simpler concepts of seasonal variation.

The series of monthly temperatures resembles a synthetic series constructed by adding a random component to a set of twelve constants where each of the constants is a component of the observations for a particular calendar month in all of the years represented by the series. Hence it seems realistic to specify that the temperature series contains a linear trend with a slope of zero, a cyclical component which is identically zero, a seasonal component which consists of the twelve unknown long run mean temperatures for each month expressed as deviations from their common mean, and a random component consisting of random selections from a normal universe whose mean is zero. All of the conventional components are included in this simple statistical specification even though the trend and cyclical components are not necessary in this particular case. On the basis of this specification, maximum likelihood estimates of the seasonal parameters can be obtained by finding the arithmetic mean of all of the available observations for each of the twelve calendar months. The twelve results are expressed as deviations from their general mean which is a maximum likelihood estimate of the single trend parameter. The series of monthly temperatures can be adjusted for seasonal variation by subtracting the estimate of the appropriate seasonal parameter from each of the observations to obtain an adjusted series having the same mean but without a seasonal component. It can also be adjusted by subtracting the estimated

normal temperature, the sum of the trend and seasonal components, to obtain a series of deviations from normal temperatures. The latter adjustment which is equivalent to subtracting from each observation the average of all available observations for the same calendar month often results in a series which is more useful in further analysis. However, both types of adjustment are in common use.

Writers on statistical methods rarely make detailed statistical specifications on which to base their analysis of economic time series. This is not surprising because manageable specifications are usually unrealistic when applied to typical problems in the study of related groups of economic series. It is difficult or perhaps impossible to formulate a statistical specification which would describe the important characteristics of a typical economic time series as well as the foregoing simple statistical specification describes the temperature series. Even if such a specification were available it is almost certain that it would be extremely difficult to devise effective methods of estimating its parameters. Hence it is worth while to examine some of the relationships of statistical specification to the problems of estimation and distribution in connection with a specification which seems to be reasonably satisfactory.

In series of monthly temperatures the trend and cyclical components which are responsible for most of the difficulties in connection with the analysis of economic time series are both represented by a single parameter. The mean annual temperature of a given locality may actually be subject to long run changes of the type described by trend functions or long cycles. Optimum statistical estimates of the parameters of functions describing such long run changes will not be obtained by means of the method described above because they are specified to be nonexistent. This is an illustration of the general principle that the application of optimum methods for statistical specifications results in estimates of no parameters except those which are specified.

Another interesting feature common to most statistical specifications is the random component. The usefulness of this device does not depend on the belief that deviations from the average temperature of a given locality are determined by means of random selections from a normal universe. It is true that many statistics have the maximum likelihood property only when the uni-

verse of the random component is exactly normal and that the exactness of the common tests of significance depends on normality of the universe from which the random component is selected. In most cases, however, moderate departures from normality have little effect on either maximum likelihood estimates or sampling distributions of statistics.

As has been previously explained, adjustment of a time series by the removal of the seasonal component is analogous to the use of first approximations in graphic regression analysis. Even when some measurable weather characteristic such as mean monthly temperature is specified to be the chief cause of variations in a monthly time series, it may be advisable to make a mechanical seasonal adjustment based on averages of some function of the series for each of the twelve calendar months. The value of the adjusted series for any given month is the observed value adjusted for mean monthly temperature which is associated on the average with that calendar month. This series can then be further adjusted for the effects of deviations from the long time average of the observed mean monthly temperatures. This method has several advantages over direct adjustment for the effects of temperature by regression methods. Its efficiency is less dependent on the form of the regression function used. There is also less danger of attributing to temperature variation the effects of customs, holidays, and characteristics of the calendar. Although the use of the method does not entirely eliminate difficulties due to the possible difference in direction of effects of high temperatures at different seasons of the year, it facilitates the investigation when such effects are present. Finally, it helps reduce the number of variables.

The Method of Monthly Means

The method of monthly means in its simplest form has been explained in connection with the series of monthly temperatures. In this section the method is extended to problems in which the trend and cyclical components are subject to important changes. It is simpler to introduce these additional components one at a time. Hart has shown that the method of monthly means results in the exact values of the seasonal components of a synthetic series when the seasonal component for each month is a constant and the cyclical component consists of the sum of complete periods of sine

and/or cosine functions of time whose periods in years are integers greater than one.¹ Hart did not consider series with random components, but the maximum likelihood property is not affected by the addition of such a cyclical component. The cyclical component need not be restricted to sums of sine and cosine functions. If the functions have no components with a period of one year; that is, if the cyclical component averages out perfectly, the method of monthly means is applicable exactly as in the case in which the cycle is composed of sine and cosine functions.

When a monthly series contains a component described by a trend function which is linear in its parameters, a seasonal component consisting of a constant value for each calendar month, and a component selected at random from a normal universe, the method of monthly means is applied by fitting the appropriate trend function by the method of least squares to annual averages using the mean deviation from trend for each calendar month as the estimated seasonal component for that month. All of the statistics have the maximum likelihood property under these conditions.

The discussion of optimum methods for determining estimates of seasonal parameters increases in difficulty when both trend and cyclical components are specified. If maximum likelihood estimates are to be derived it is necessary that the cycle does not affect the determination of the trend coefficients. In the case of a polynomial trend of degree k , this implies that the sum of products of the cyclical component multiplied by corresponding values of each of k powers of time deviations from the midpoint of the series must be equal to zero. In addition the cyclical component must average out perfectly when deviations from trend are averaged by calendar months. It is interesting to note that the synthetic series used by Hart to illustrate the method of monthly means has a downward trend due to the effects of the cycle. If he had used deviations from trend as his discussion suggests, the estimates of the seasonal components for the first six months would have been lower than the numbers which he added to the smooth cyclical component in the construction of the series while the estimates for the other six months would have been too

¹William L. Hart, "The Method of Monthly Means for Determination of a Seasonal Variation," Journal of the American Statistical Association, XVIII (September, 1922), 341-49.

large even in the absence of a random component.

The effectiveness of the method of monthly means is greatly reduced by important discrepancies from the types of cyclical fluctuations which have been specified. Hence, the simple averaging process advocated by Hart should be supplemented when the cyclical component is irregular. The device which is most frequently used to eliminate approximately the effects of irregularity in the cyclical component is a centered moving average with a period of twelve months.

A moving average in which observations for each of the twelve calendar months receives equal weight eliminates a seasonal component consisting of twelve constant values. In the absence of a cyclical component such a moving average provides a good estimate of the trend component since its deviations from such a trend are wholly due to variations of the random component on which the moving average acts as a smoothing formula. In fact, the variances of these deviations are one-twelfth the variance of the random component. When a cyclical component is included in the specification, the moving average becomes an estimate of trend plus cycle with a bias in the direction of concavity which is slight for the case of a smooth long-period cyclical component. It should be noted that deviations from such a moving average are composed chiefly of seasonal and random components under these conditions even when the estimates obtained by the unmodified method of monthly means are invalidated by the effects of cyclical fluctuations which fail to average out. Hence, deviations from a centered twelve months moving average can be averaged by months to obtain estimates of the seasonal parameters which would have the same properties as similar estimates derived from deviations from trend in the absence of cycles if it were not for the effects of bias in the direction of concavity. As suggested in the preceding chapter such bias which is also influenced by curvilinearity of the trend component can be partially avoided by the use of weighted averages designed for the purpose or by graphical smoothing.

The restriction of the cyclical component to smooth long-period cycles is unrealistic for series subject to abrupt changes of a cyclical nature. Such series are usually analyzed as if deviations of the cyclical component from regularity were a part of the random component. Under these conditions the estimate of each

seasonal parameter is an average of a distribution which is apt to be leptokurtic to a pronounced degree. In order to provide for such a set of problem conditions it is necessary to specify that the universe from which the random component is selected is leptokurtic and to use methods of averaging which are appropriate for such universes. The median which is a maximum likelihood average for double exponential universes is often used. The modified median or the mean of central items is another average which is strongly recommended for leptokurtic universes. Both averages are in common use in the statistical estimation of seasonal parameters.

The method of monthly means is seldom used in the adjustment of economic time series for seasonal variation because it is rarely appropriate to specify that the seasonal component is a set of twelve constants. The seasonal component of most time series is described more accurately as a function of the other components. For example, it may be proportional to the trend or to the trend plus cycle. The seasonal parameters are therefore specified to be factors of proportionality. With the modifications of the method of monthly means herein suggested the method would compare favorably in both accuracy and realism with other familiar methods in common use if it were applied to logarithms instead of to the original series.

The Method of Trend Ratios

The method of trend ratios as presented by Hall is similar to the method of monthly means applied to the ratios of the original observations to the corresponding ordinates of a least squares trend fitted to the series.² The arithmetic mean of trend ratios for each calendar month is computed. Each of these means is divided by the general mean to obtain an index of seasonal variation from the corresponding trend ratio.

It is difficult to formulate a statistical specification on the basis of which the method of trend ratios is the optimum method of estimating the seasonal parameters. Since the least squares trend is determined first, it is desirable that deviations from trend should be approximately homoscedastic. Later

²Lincoln W. Hall, "Seasonal Variation as a Relative of Trend," Journal of the American Statistical Association, XIX (June, 1924), 156-66.

these same deviations are averaged out in relative residual form when the trend ratios are averaged to find indexes of seasonal variation. Hence, it is desirable that relative deviations should also be approximately homoscedastic. In the case of an increasing trend function this implies that both the variance of the random component and the amplitude of the cyclical component are increasing functions of time. Both absolute and relative deviations from trend cannot be approximately homoscedastic if the trend increment is important. On the other hand if the trend is specified to be constant, the seasonal component, which is proportional to trend, is a set of twelve constants. Under these conditions, some modification of the method of monthly means would be superior to the method of trend ratios.

The method of trend ratios depends on the process of averaging to eliminate the effects of deviations from regularity in the cyclical component. Because of the effects of unusual deviations on the arithmetic mean, Falkner strongly recommends the use of an average of central items in the derivation of the indexes of seasonal variation.³ The geometric mean, a maximum likelihood average for logarithmic normal universes, is often used in the averaging process because ratios have a lower limit but no upper limit. Hence, the skewed distribution of a hypothetical universe of ratios might be approximated by a logarithmic normal frequency function. The average of central items which was discussed in greater detail in the preceding section is probably to be preferred to other averages in most cases.

The device of moving averages could be used to increase the effectiveness of the method of trend ratios. When the cycle is irregular, as is usually the case, the seasonal parameters can be estimated more accurately if a centered twelve months moving average of deviations of trend ratios from their general mean is subtracted from the trend ratios before the averaging process is applied. In most cases this reduces the undesirable effects of cyclical fluctuations which do not average out when the usual method of trend ratios is used.

³Helen D. Falkner, "The Measurement of Seasonal Variation," Journal of The American Statistical Association, XIX (June, 1924), 167-79.

The Ratio-to-Moving-Average Method

When the seasonal component is proportional to trend plus cycle the appropriate method for adjusting the series for seasonal variation is to divide each observation by the appropriate index of seasonal variation. There are two methods which seem appropriate for estimation of seasonal parameters for such a series. One of these is the method of monthly means applied to the deviations from a centered twelve months moving average of the logarithms of the original series. The usual adjustment can be applied to the antilogarithms in order that final seasonal indexes average 100 in percentage form. The additional device of using the mean of central items to estimate the logarithms of the seasonal indexes is also applicable.

The other method which is more commonly used is to average ratios to a centered twelve months moving average to obtain the first approximations to the indexes of seasonal variation which are then adjusted by dividing by the general mean to obtain indexes which average 100 when expressed in percentage form. There seems to be little basis for preferring one method to the other for problems in which either would be appropriate. Even the labor of looking up logarithms is offset by the substitution of addition and subtraction for the computation of ratios. The ratio-to-moving-average method is probably the most efficient of the familiar methods in common use. Its chief advantage over methods in which moving averages are not used is its direct measurement and elimination of irregularities of the cycle. Other well-known methods depend upon the process of averaging to eliminate such irregularities.

The ratio-to-moving-average method is most effective when the seasonal component is proportional to trend plus cycle, but it is superior also when the seasonal component is proportional to trend unless other methods are modified by the use of an appropriate moving average. For example, when the seasonal component is proportional to trend the seasonal indexes can be estimated by averaging ratios whose denominators are trend ordinates. The numerators of these ratios should be sums of the same trend ordinates and the corresponding seasonal components. These sums are approximated by a centered twelve months moving average. This method eliminates statistically the cyclical component from both numerators and denominators of the ratios. In the unmodified

ratio-to-trend method the cyclical component is eliminated from the denominators only. On the other hand, the cycle appears in both numerators and denominators in the ratio-to-moving-average method. Since adding a small constant to both terms of a fraction whose value is not far from unity changes the value of the fraction less than adding the same constant to the numerator only, it is apparent that the appropriate ratios are approximated more closely by ratios to moving average than by ratios to trend.

The method of adjustment of a series for seasonal variation depends not only on the form of the relationship of the four components but also on the desired form of the adjusted series. When the seasonal component is proportional to trend plus cycle, the seasonal component is eliminated from the original series by dividing each observation by the appropriate index of seasonal variation. The adjusted series is often expressed as ratios to trend, deviations from trend, or cyclical relative residuals. Products of trend ordinates and indexes of seasonal variation are sometimes subtracted from the original series to obtain seasonally adjusted deviations from trend. Another form often used is obtained by reducing either relative or absolute deviations from trend to standard deviation units.

The Link Relative Method

The method of link relatives has advantages over other familiar methods when the series is subject to sharp cyclical fluctuations and/or other changes which are practically equivalent to discontinuities in the series. Since it is based on ratios to preceding observations each discontinuity necessitates the loss of only one of the link relatives. Under these conditions it is especially important to use the median or the average of central items. If the median is used, results identical with those obtained by the usual method as first described by Persons⁴ can be obtained from the series of logarithmic first differences by a method which is easier to explain. A median of the logarithmic first differences is determined for each month. Beginning with February the medians are added cumulatively and the subtotals are adjusted by subtracting one-twelfth of the final total from the

⁴Warren M. Persons, "Indices of Business Conditions," The Review of Economic Statistics, Preliminary Volume I (1919), 5-107.

first subtotal, two-twelfths of the final total from the second subtotal, and so on until the adjustment of the January median reduces it to zero. These adjusted subtotals correspond to the logarithms of the adjusted chain relatives of the method as usually described. Adjusted chain relatives are obtained by taking antilogarithms which are then reduced to seasonal indexes which average 100 in percentage form. The method of adjusting the series for seasonal variation is the same as for the ratio-to-moving-average method.

The adjustment ordinarily applied to chain relatives is simple in logarithmic form. It eliminates the bias contributed by an exponential trend which is linear in terms of logarithms. The bias caused by other types of trend is approximately eliminated also if the relative rate of change is not subject to great changes. In addition, it provides an adjustment for bias introduced in the averaging process.

In the link relative method cyclical difficulties are limited to the effects of single month to month changes in the cyclical component on single link relatives, but unfortunately, these constitute a bias in the positive direction for all upward phases of the cycle and a negative bias for all downward phases. Since the median is usually used for averaging link relatives, unchanging phases of the cycle have too much weight in determining the seasonal indexes. This difficulty can be avoided by differencing the logarithmic series twice instead of once before applying the averaging process.

The Logarithmic Second Difference Method

A method based on logarithmic second differences, developed by Theodore O. Yntema and the writer, is closely related to the method which has been described in published articles as the variate difference method.⁵ The method, which has not been previously discussed in terms of logarithms, involves the following steps:

1. Compute second central differences of the logarithms of the original series.

⁵Richard A. Robb, "Variate Difference Method of Seasonal Variation," Journal of the American Statistical Association, XXIV (September, 1929), 250-57; and Victor S. von Szeliski, "Comment of 'The Variate Difference Method of Seasonal Variation,'" Journal of the American Statistical Association, XXV (March, 1930), 83-84.

2. Find the average second difference for each month using any satisfactory method of averaging.
3. From each of the twelve averages subtract their mean.
4. Add the adjusted averages subtotalling after each item.
5. From each subtotal in step 4 subtract the mean of the subtotals.
6. Add the adjusted subtotals in step 5 cumulatively.
7. Obtain antilogarithms of the subtotals in step 6.
8. Divide each antilogarithm by the mean antilogarithm to obtain indexes of seasonal variation which apply to months following those to which the corresponding central differences refer.

The use of central differences is recommended because each central difference is equal to twice the negative of the logarithm of the ratio of the corresponding observation to the geometric mean of adjacent observations. This property is often useful especially in case of changing seasonal pattern.

When the logarithms of a long series contain a random component, the sampling variance of mean central differences derived by the logarithmic second difference method is six times the variance of the monthly means of logarithms. This fact has been used as an argument tending to show the relatively large effect of the random component on statistics based on finite differences.⁶ Such an argument is misleading because the ratio of sampling variances given above is irrelevant so far as effects on the accuracy of estimates of seasonal indexes is concerned.

It should be understood that practical reasons for the use of logarithmic differences and averages other than the arithmetic mean have no relation to components which are really random selections from a normal universe. These devices are useful only when the series differs in some important respect from a simple synthetic series. In order to compare alternative methods on the basis of sampling errors alone it is imperative that all such difficulties be abstracted from. It can then be shown that sampling variances of estimates of the same parameter derived by the three methods considered are almost identical because differences among the three sets of estimates are simple functions of sampling errors of one

⁶Robb, op. cit., p. 252.

or two observations which could not be important in a fairly long series.

Consider a series of antilogarithms of 120 random selections from a normal universe with zero mean multiplied by a set of monthly indexes of seasonal variation. Maximum likelihood estimates of the logarithms of the seasonal indexes can be obtained by applying the simple method of monthly means to the logarithms of the series. The sampling variance of each of these statistics is one-tenth of the variance of the normal universe from which the random logarithms are selected. For every sample described by the foregoing specification, the two sets of estimates of the logarithms of the indexes of seasonal variation derived by the link relative method in mean logarithmic first difference form and by the mean logarithmic second difference method differ from the set obtained by the method of monthly means of logarithms by a function of the sampling errors in the first and the last observations.

This is apparent upon consideration of the fact that complete agreement among the three sets of estimates can be secured by arbitrarily adding an observation at the beginning of the series equal to the last observation of the original series when logarithmic first differences are used. The logarithmic second difference method requires in addition an observation at the end which is equal to the first observation of the original series.

Adjustment for Changing Seasonal Variation

When the components of a time series can be described satisfactorily by any of a group of simple specifications, some of the methods already suggested can be used to eliminate the seasonal component. It should be noted, however, that the parameters of relationship between the seasonal component and the other components are a set of constants in every specification thus far considered. When a series corresponds to this type of specification its seasonal pattern is said to be constant.

There are many types of change to which the seasonal parameters may be subject. The simplest type of change in seasonal pattern, and perhaps the most useful one for purposes of specification, is one in which some of the parameters are subject to a gradual change which can be represented by a smooth function of time, that is, some seasonal parameters have trends. It is doubtful if this

specification should be used when variations in the effects of weather factors are due to a succession of abnormal weather conditions rather than to changes related to the series. It is better to use regression methods in such cases because mechanical methods of adjustment for seasonal variation are not well adapted for problems in which the effects of deviations from average weather conditions are specified to be a part of the seasonal component.

Mendershausen has reviewed the work of previous writers on the problem of eliminating changing seasonal variation.⁷ Since it seems unnecessary to repeat Mendershausen's excellent and detailed critical discussion, only the most important aspects of the problem will be discussed here.

When a series is subject to changing seasonal variation, it is especially important to base estimates of seasonal parameters on ratios or other functions which are as free as possible from the cyclical component not only on the average but also individually. Methods in which the device of moving averages is used to eliminate the effects of irregularities of the cycle seem to be most effective for the measurement of changing seasonal variation although logarithmic second differences might be used without adjustment in case the cycle is smooth. Since it has been shown that the device of moving averages can be used in connection with most of the common methods of measuring seasonal variation, any of the common methods can be modified so as to provide a set of seasonal indexes for each year instead of one set to be used for the entire series. The simplest way to determine a set of parameters for a series with a changing seasonal pattern is to fit trends by any acceptable method to the twelve sets of functions which would be averaged in the case of a constant seasonal pattern. The set of twelve trend values for each calendar year is used in the same way that the twelve averages of central items would be used in the simple case to compute indexes of seasonal variation for that year. For example, in the case of the ratio-to-moving-average method the trend value of the January ratio to moving average for a given year can be used as a seasonal index for January of that year. It is customary, however, to adjust

⁷Horst Mendershausen, "Methods of Computing and Eliminating Changing Seasonal Fluctuation," Econometrica, V (July, 1937), 234-62.

these to average 100 in percentage form as is usually done for the case of constant seasonal pattern. In some cases "years" other than calendar years would be more useful in order that breaks would come at an appropriate season.

Another type of changing seasonal variation is a change in the amplitude of the seasonal pattern. The deviations of the various seasonal indexes from unity are proportional to a function of time which is subject to smooth changes over the cycle or to changes in one direction due to technological progress. According to the method proposed by Kuznets⁸ for this type of changing seasonal pattern, one set of values for the deviations of seasonal indexes from unity is derived by the ratio-to-moving-average method in the same way as for a constant seasonal pattern. The variable factor of proportionality is computed as a linear function of the original ratios to moving average. Let r represent deviations from unity of the set of seasonal indexes computed as for the case of constant seasonal pattern. Let x represent deviations from unity of ratios to moving average. The variable factor of proportionality is a linear function of twelve x 's and it applies to the sixth or seventh month of the time period. The coefficient of a particular x is the corresponding r divided by the sum of squares of the r 's. It seems an unnecessary refinement to require that the linear function of the x 's be centered exactly, but the use of the same value of the function for all months of each calendar year leads to discontinuities in the pattern of seasonal indexes between each December and the following January.

The adjustment of time series for seasonal variation on the basis of the specification that the amplitude of the seasonal pattern is changing does not seem to be of important general applicability. It seems worth while to mention the method, however, because cases might arise in which the method would be useful.

Graphic Methods of Adjustment for Seasonal Variation

The work of previous writers on the use of graphic methods of measuring and eliminating seasonal variation has been reviewed

⁸Simon Kuznets, "Seasonal Pattern and Seasonal Amplitude: Measurement of Their Short-Time Variations," Journal of the American Statistical Association, XXVII (March, 1932), 9-20.

by Spurr.⁹ As presented by Spurr the graphic method is flexible and can be adapted so as to be appropriate for a wide variety of specifications including changing seasonal variation. The advantages of the graphic method can be secured without appreciable sacrifice of the advantages of any of the computational methods in common use. In fact, graphic techniques can be supported by moving averages or any other computation device which may seem desirable. It is possible, however, to adjust a series for seasonal variation without making any computations. In addition to the great saving in time and effort which can often be made by the use of the graphic method it is of theoretical interest to mention that it is often possible to smooth a moving average graphically in such a way that it represents the trend plus cycle more effectively.

Miscellaneous Problems in Seasonal Adjustment

A special type of seasonal adjustment is necessary when a series is affected by the varying lengths of months and it is not desirable to include this adjustment with the ordinary mechanical adjustment for seasonal variation or when no other adjustment is to be made. In such cases monthly figures are often divided by the number of days in the month to reduce the series to an average per day basis. This adjustment seems rather obvious and, in some cases, there could be little argument as to the number of days to use. It is not always clear, however, whether it would be more useful for purposes of analysis to divide by the number of calendar days or the number of working days. The following description of a method of adjustment used in the Bureau of Foreign and Domestic Commerce explains the reduction of grocery chain store sales to average per weighted working day basis.

Not all months have an equal number of working days, nor are sales of grocery chains evenly distributed throughout the week. According to information supplied by some of the contributing firms, nearly 40% of the weekly business of chain grocery stores is normally done on Saturday. For the purpose of converting the sales figures to an average daily sales basis per calendar month, the number of working days in each month has been calculated for the years 1929 to 1937 inclusive (table 3). The days of the week were evaluated from the sales

⁹William A. Spurr, "A Graphic Method of Measuring Seasonal Variation," Journal of the American Statistical Association, XXXII (June, 1937), 281-89.

experience of the firms and the following weights assigned: Monday, 0.6; Tuesday, 0.7; Wednesday, 0.7; Thursday, 0.7; Friday, 1.0; Saturday, 2.3; total, 6.0. The number of working days in any month was then calculated by adding the weights of all the business days in that month. No allowance was made for holidays, except as noted below, it being assumed that food for consumption on holidays is bought on the preceding day and that the total sales are not affected. However, when a holiday falls on the first working day of a month, six-tenths of a day, the weight of Monday, is subtracted from that month and added to the preceding month. The sales figures on a calendar-month basis were then divided by the calculated number of working days in the respective months to arrive at the average daily sales for the different months.¹⁰

One of the interesting features in the method described in the foregoing quotation is the set of weights which was assigned to the days of the week. These correspond to seasonal indexes based on the intraweekly pattern of sales and were probably derived from daily series by a method analogous to those previously discussed for use with monthly series. It is not necessary to compute these indexes accurately because they affect only those days in excess of four weeks, and the effect of a reasonably small error in the weights is negligible.

Another component of the method which is worthy of notice is the combination of two types of adjustment in one process. Whether this combination of an empirically derived seasonal adjustment for the intraweekly pattern in grocery sales with the ordinary adjustment to average per day basis is the most satisfactory method of adjustment depends on the relative usefulness of series adjusted by this, as contrasted with the best alternative method, for purposes of comparison with other series. Simplicity and the probable degree of permanence of the intraweekly pattern of weekly sales should be considered. Thus, choice of a method of seasonal adjustment depends on the purposes for which the adjusted series is to be used.

There are many problems in which the adjustment for holidays is not so simple as in the foregoing example. Holidays and half holidays present many difficulties which can hardly be dealt with by means of a general method. It is usually necessary to use a considerable amount of ingenuity in order to devise methods of dealing with the problem and these are seldom of general appli-

¹⁰Reba L. Osborne, "Index of Grocery Chain Store Sales," Survey of Current Business, May, 1937, pp. 15-16.

cation. An interesting example is the adjustment of department store sales for the effects of the changing date of Easter.¹¹ The solution in this case consisted of an adjustment derived by regression analysis which was applied to the seasonal indexes for March and April. The adjustment served to divide the large total of the indexes for the two months on the basis of the date of Easter for each year.

Many attempts have been made to avoid the difficulties due to the peculiar characteristics of the various calendar months and the customs associated with them. One of these is the use of the thirteen-month calendar for the purpose of simplifying internal records in order to facilitate administrative control. Difficulties similar to those for which adjustment was made in the chain store sales index are avoided by the use of such a calendar because each month contains exactly four weeks. A new difficulty is introduced, however, by the fact that the "year" contains only 364 days. Another method for avoiding calendar problems is to adopt a business month which begins on some day other than the first of the corresponding calendar month. For example, corporations having many accounts on which monthly payments are to be made may begin new "months" on the twenty-first instead of the first of each calendar month. Many delinquencies as of the close of February and thirty day months occur because payments on accounts opened on the thirty-first of long months are apt to be postponed until the first of months following a short month.

Other problems which require the application of techniques similar to those previously described include the adjustment of weekly series, monthly series based on a thirteen-month calendar, and series reported at intervals shorter than one day; such as, shares traded on stock exchanges, density of traffic, frequency of telephone calls, use of electric power, etc. When adjustment of such series is necessary, problems which arise are mostly related to the type of specification which describes most adequately the special characteristics of individual problems. These are not often of general interest.

¹¹Aryness Joy and Woodlief Thomas, "Adjustments for the Influence of Easter in Department Stores Sales," Journal of the American Statistical Association, XXII (December, 1927), 493-96.

Advantages of Seasonal Adjustment

It is necessary to consider problem conditions carefully in order to decide whether it is appropriate to apply mechanical methods of seasonal adjustment. Similar seasonal patterns in two series to be compared may occur because weather affects both series in a similar manner. The relation between seasonal patterns of egg production and ice cream sales is thus largely attributable to effects of a common cause. On the other hand, the seasonal component of one of the series may be caused by weather and institutional factors while the seasonal variation in the first series is one of the causes of seasonal variation in the second series. An example of two series related in this way is automobile production and steel production. It is not possible to devise a statistical technique which will distinguish between these two types of relationship on the basis of the observed series alone. Like all other partial relationships derived by statistical reduction of observed series, relationships between seasonally adjusted series may or may not represent important net relationships between the economic variables represented by the series. This principle is especially important in regression analysis of time series.

The fundamental reason for adjusting a series for seasonal variation is to facilitate comparison with other series. The seasonal adjustment may, in most cases, be considered as an attempt to eliminate the effects of weather factors and/or institutional factors associated in some way with the weather or the calendar. In some problems, the effects of weather factors can be eliminated more satisfactorily by means of regression functions as in Bard's study of steam consumption which was explained briefly in the first chapter. While the use of regression functions has many advantages, there are several reasons for the use of mechanical methods for seasonal adjustment. Some of these are considered below.

There is no satisfactory means of classifying and measuring some institutional factors or the variable characteristics of the weather which serve as a partial explanation of them.

It is difficult to specify useful functional forms for regressions of observed series on measurable institutional and weather factors. In other words, the scales of measurement for these factors are often unsatisfactory for regression analysis even in the case of measurable factors.

A multiple regression analysis in which the two or more series to be compared and the various weather and institutional factors are included as variables would be appropriate under certain specifications. Such a method would often be inapplicable even if all factors were measurable because the effect of a given factor is a function of the expected value of other factors for that month as well as its own expected value. This familiar characteristic of weather and institutional factors necessitates joint regression analysis and increases the number of variables to such an extent that even graphic techniques would not be feasible. For such problems it is usually more useful to consider the elimination of the typical seasonal component as the first stage of a regression analysis in which the seasonal adjustment is an approximate correction for the effects of average values of weather and institutional factors. The effects of deviations of measurable factors from their average values should be removed by ordinary regression analysis as suggested previously. This method minimizes the ill effects of choice of inappropriate functions for the regression of observed series on weather factors which might be serious if the typical seasonal component were not first removed.

The use of a mechanical method for the elimination of seasonal variation has another advantage over the method of least squares multiple regression analysis, even in problems in which the latter method might be used. This advantage results from the common use of the average of central items in the determination of seasonal indexes to obtain more reliable averages from leptokurtic distributions. This is especially valuable when unusual deviations are caused by a variable whose fluctuations have been of major importance so infrequently that little or no statistical evidence exists as to the form of the net regression function of the observed series on the causal variable.

Sometimes it is desirable to apply the seasonal adjustment in the process of deriving the statistical relationship among several variables by successive approximations. An excellent sample can be found in a study of department store cost functions.¹² The monthly cost series for a particular department was adjusted

¹²Joel Dean, "Department Store Cost Functions," an essay to be published in the Henry Schultz Memorial Volume.

for variations in the advertising rates and all allocated general store overhead costs were excluded. The net regression of the resulting cost series on the two most important independent variables—number of transactions and average value of transaction—was then determined by the Ezekiel-Bean graphic method. Residuals from this regression were used to estimate the seasonal component by the method of monthly means. The appropriate seasonal component was then subtracted from the cost series used in the graphic studies to obtain the series finally used in a least squares regression analysis. The sequence of methods used is of interest—mechanical statistical deflation, graphic regression analysis, seasonal adjustment, and least squares regression analysis.

CHAPTER IV

ECONOMIC RELATIONSHIPS AND REGRESSION ANALYSIS

Abstraction from Important Problems

Regression analysis is used extensively in the statistical determination of the functional forms and the parameters of economic relationships. The difficulties encountered in econometric studies are so numerous and interrelated that it is necessary to concentrate attention on a limited number for purposes of exposition. For this reason the effects of secular and cyclical changes are abstracted from as much as possible in this chapter. Abstraction from characteristics peculiar to time series serves to emphasize the types of problems which are superimposed upon those which arise because of important dynamic changes in the economic system.

Definition of Net Relationships in Regression Analysis

In many studies of related economic series, emphasis is placed on the estimate of a single net relationship. The most familiar example is the net relationship between the rate of sale or consumption of a given commodity and its price in a given market. This implies that other variables are to be held constant, but it is not easy to define exactly what is meant by holding other variables constant and which variables are assumed to be held constant when a given net relationship is being discussed.

The net relationship between the dependent variable and one of the independent variables is obtained by assigning constant values to all of the other independent variables. The mathematical form of the relationship is the same regardless of the values assigned to any of the variables which are held constant. There is no statistical evidence that the relationship holds for values assigned to any of the variables outside the region in which simultaneous observations on all of the variables are available.

The shape of a net relationship may depend upon the particular values at which the other variables are held constant.

Although it is important to recognize this possibility, it is unnecessary to discuss the subject in detail. For the case in which reliable observations are available over the range in which attention is directed, methods of dealing with such relationships are familiar to students of regression analysis in connection with the Ezekiel-Bean method of graphic regression analysis.¹

Some statistical economists limit the use of the graphic method to graphic scaffolding in exploratory work and rely exclusively on mathematical methods for the final determination of the empirical relationship. This is due in part to difficulties in evaluating the reliability of regression functions obtained by the graphic method. It is true that the reliability of the graphic regression function of Y on X depends largely on the number of observations which represent Y -arrays close to the one for which the reliability is to be evaluated even when it is known a priori that the sample is selected at random from normal homoscedastic arrays of Y and that the regression function is smooth. It does not seem to be generally known that the reliability of mathematical regression function suggested by graphic scaffolding is subject to the same limitation. The applicability of the familiar standard error of the function depends on a priori knowledge of the functional form of the regression function. Exclusion of small values of test criteria caused by making specifications based on characteristics of the sample arises just as truly in the early stages of an investigation as it does in the later stages. For this reason graphic methods should be applied whenever they are useful even to the extent of supplementing a mathematical analysis with a graphic analysis of the residuals. When the study warrants the effort, graphic scaffolding should be followed by a method like least squares which is independent of the order of consideration of the variables. This mathematical analysis should be followed by a study of the residuals by graphic methods which are effective at this stage in suggesting changes which would make the appropriate allowances for possible inadequacies of the mathematical function used or for the effects of extreme deviations from the regression which often distort regression functions fitted by mathematical methods. This is especially true when specifications for which optimum estimates

¹Mordecai Ezekiel, Methods of Correlation Analysis (New York: John Wiley and Sons, 1930).

can be derived by least squares are unrealistic. Of course, the process suggested should be used in its entirety only when the determination of the appropriate form of the regression function is of sufficient importance. In such problems, residuals should be plotted against each of the variables in turn and the regression function of the residuals on each variable should be added to the net relationship of that variable with the dependent variable unless the residual regressions are sufficiently important or intercorrelations among independent variables sufficiently high to justify successive approximations by graphic regression analysis using the residuals as the dependent variable.

It is also important to consider the possibility that the form of a regression function may differ markedly from that suggested by the relationship within the region which contains simultaneous observations on all of the variables. For example, it is doubtful whether the functional form of most statistical demand relationships are sufficiently dependable to permit extrapolations. It is likely that many related commodities act more like close substitutes at unusual price differentials than at more typical ones. Potential substitution becomes effective in many cases only after the price differential exceeds certain minimum costs of adjustment. Sometimes transportation costs must be incurred in order to make large scale substitution possible. Standard errors which measure the unreliability of extrapolations under favorable conditions are not applicable in this type of problem because they depend on the specification that the functional form of the function is known to be appropriate for all ranges of values of all of the variables and that effects of variables not included in the analysis are random. When the prices of close substitutes are included in a demand study, the net demand relationship for any one of the commodities is unreliable. At best it can apply to an extremely limited range when prices of related commodities are held constant. Holding such prices constant at unusual price differentials is untenable.

That extrapolation is not dependable is a principle that applies also to variables not included in the analysis. Schultz applied this principle in connection with unusual price ratios of oats and wheat,² but he did not emphasize it in his discussion of

²Henry Schultz, The Theory and Measurement of Demand (Chicago: The University of Chicago Press, 1938), p. 462.

the interpretation of statistical demand curves.³ In a search for an example with which to illustrate this principle, the writer made a simple graphical analysis of Schultz's data for oats and corn, two closely related commodities. Schultz's studies of the related demands for these two grains indicated that the price of oats was relatively unimportant in the demand for corn and vice versa.⁴

The writer's study of the related demands of corn and oats was carried out in terms of quantity ratios and price ratios of the two grains. A scatter diagram on double logarithmic scale indicates a linear regression in which the residuals are relatively small. This simple preliminary analysis supports the contention that a demand function for corn in which the price of oats is omitted can be applicable only for typical price ratios of corn and oats.⁵ Other related prices which may be found to be statistically insignificant affect the demand function for corn in a similar manner.

The Mathematical Theory of Exchange

Statistical derivation of demand curves is an important application of regression analysis in economic research. Hence, it seems appropriate to summarize briefly the mathematical theory of exchange. The relevant portion of the theoretical discussion with which Schultz prefaced his statistical demand studies⁶ is supplemented by material of interest taken from a more extensive treatment of related theories by Hicks.⁷

Consider a group of individuals in a market. Each has a given amount of each of a number of commodities. On the basis of known preference systems, each individual tries to secure by exchange that combination of commodities which is most preferable to him within the limits of his resources or wealth. All exchanges are made with the privilege of recontract and the indifference functions are all continuous and convex towards the origin at all points. The amounts of the various commodities which will be offered in exchange by their original owners and all of the exchange

³Ibid., pp. 71-72.

⁴Ibid., p. 594.

⁵Ibid., p. 276.

⁶Ibid., chap. 1.

⁷John R. Hicks, Value and Capital (Oxford: The Clarendon Press, 1938).

ratios can be expressed in terms of solutions of a set of simultaneous equations.

The condition that the indifference functions are convex implies that, at any set of exchange ratios, any individual is able to choose from among the infinite combinations of commodities which have a market value equal to his wealth a single combination which is preferable to all others. As Hicks put it, ". . .there is a sufficient degree of regularity in the system of wants (and also, as we shall see later, in the productive system) for any set of quantities in the neighbourhood of those with which we are concerned to be a possible position of equilibrium at some system of prices."⁸ Another way of expressing the same thing is to state that "the marginal rate of substitution diminishes in every direction at every position with which we shall be concerned in our analysis."⁹

The privilege of recontract implies that all exchanges are made on the basis of ratios established at equilibrium. Insisting that recontract is unrealistic, Hicks shows that "a change in price in midst of trading has the same sort of effect as redistribution of wealth."¹⁰ Hicks concludes on the basis of an analysis of such "income effects" that one should ordinarily expect stable equilibrium and that the presence or absence of recontract usually has little effect on the prices which are finally established.

The specification of continuity for individual indifference functions and the demand functions which are derived from them is unrealistic. Hicks says:

This convenient assumption of continuity does, of course, always falsify the situation a little (or sometimes more than a little) as far as the individual consumer is concerned. But if our study of the individual consumer is only a step towards the study of a group of consumers on the market, these falsifications can be trusted to disappear when the individual demands are aggregated.¹¹

The theory which describes static equilibrium in terms of the original quantities of many commodities and their distribution among many individuals would be of little value if it were not for the fact that the analogous solution for a simple dynamic

⁸Ibid., p. 23.

⁹Ibid., p. 25.

¹⁰Ibid., pp. 127-29.

¹¹Ibid., p. 11, n. 4.

model, the stationary state, is of the same form. For example, Schultz says:

If instead of thinking of these equations as relating to a single exchange, we think of them . . . as describing the average of many similar situations, each extending over sufficiently long periods, but not involving secular changes in the utility function, then . . . [the expressions corresponding to wealth in the static case] is equal to the income of the individual in the average or typical period.¹²

Hicks says:

The stationary state is that special case of a dynamic system where tastes, technique, and resources remain constant through time. We can reasonably assume that experience of these constant conditions will lead entrepreneurs to expect their continuance; so that it is not necessary to distinguish between price-expectations and current prices, for they are all the same. We can assume, too, that entrepreneurs did expect in the past that to-day's prices would be what they now turn out to be; so that the supplies of commodities are fully adjusted to their prices.¹³

Since a stationary state is always in equilibrium, it is not possible to derive statistical demand functions because observations on prices and quantities at successive points in time do not show the necessary variation. However, statistical demand curves can be derived under conditions not greatly different from those of the stationary state where there are small unexpected changes in supply from period to period.

Small Unexpected Changes in Quantities

Of the large number of commodities in a stationary state, consider three unimportant commodities, A, B, and C where B and C are extremely close substitutes while no other relationship of substitution is close. Let the system be changed in such a way that the quantities of these three commodities are subject to small unexpected variations from period to period such as effects of minor variations in weather on crop yields. Under these conditions the method of multiple regression analysis of time series might be used to determine statistically the parameters of the demand functions for the three commodities. These functions are specified to be linear in terms of logarithms for purposes of

¹²Schultz, The Theory and Measurement of Demand, p. 30.

¹³Hicks, op. cit., p. 117.

simplicity and all series are based on complete and accurate data.

It is interesting to consider the simple regression of the price of A on the quantity of A in terms of logarithms. For sets of periods in which the quantity of A is the same, prices of A differ because of shifts in the prices of B and C. If logarithms of ratios of observed to expected quantities of the three commodities are independent and normally distributed, the simple regression coefficient of the logarithm of the price of A on the logarithm of the quantity of A is (except for neglected income effects) a maximum likelihood estimate of price flexibility of A. The maximum likelihood property does not depend on random "errors" in the variables among which the statistical relationship is being determined because all series are based on complete and accurate data by hypothesis. Deviations from the regression are the effects of neglected variables some of which should in this case be used as independent variables to increase the accuracy of the estimates. This should not be construed to mean that price should always be used as the dependent variable; models can be constructed in which the regression coefficient of the logarithm of quantity on the logarithm of price is a maximum likelihood estimate of the elasticity of demand.

Even when it is appropriate to determine a single parameter of the demand relationship for a commodity in the form of a simple regression coefficient, it is desirable to determine as many as possible of the other parameters. Since income effects involve all individual incomes, it is impossible to deal with the large number of series necessary for a complete analysis. Under the conditions of the example it is unlikely that it would be possible to determine many of the coefficients. If A, B, and C are the only commodities whose prices are affected appreciably by changes in the supply of these three commodities it would be useless to include the prices of other commodities as variables in a statistical demand function, but logarithms of the prices of the three commodities should be considered and it might be useful to include also an index number designed to measure some relevant characteristic of the income distribution. For present purposes it is well to concentrate attention on the estimation of the parameters of the demand for each of the commodities as a function of the three prices.

Consider the problem of estimating the parameters of the linear relationship which expresses the logarithm of the quantity of A as a linear function of the logarithms of the prices of A, B, and C. If the relationship were perfect four observations would ordinarily suffice to determine exactly the four parameters of the relationship. In view of neglected income effects which prevent an exact determination of the parameters, estimates based on such a small number of observations would not be statistically reliable. If B and C are sufficiently close substitutes, it may be advisable, for some purposes, to combine them into one composite commodity. Under these conditions their prices have an extremely high correlation with each other. Since the standard error of the regression coefficient on any independent variable is an increasing function of the multiple correlation coefficient of that independent variable on the others,¹⁴ the partial regression coefficient of the demand for A on the composite price of B and C would be more reliable than the partial regression coefficient of the demand for A on the price of either B or C with the price of the other held constant. The use of composite commodities is particularly useful for extrapolation into ranges of unusual price ratios of B and C or for predicting the effect on the demand for A of a change in the price of either B or C when the price of the second commodity does not change in the usual manner with respect to the price of the first.

Another principle can be explained in connection with a statistical study of the demand for B in which the price of C is not taken into consideration. This omission may avoid the difficulties of high intercorrelations, but the estimate of the absolute value of the elasticity of demand is practically double that of the two commodities considered as a composite if the two commodities are extremely close substitutes and of about the same importance in the economy. This is apparent on consideration of the limiting case in which the commodities are identical but one-half of the average output is designated C and held constant while the remainder is subject to small changes in quantity from period to period. Since the prices are always identical and the elasticity

¹⁴John H. Smith, Tests of Significance: What They Mean and How To Use Them (Chicago: The University of Chicago Press, 1939), p. 31 and p. 85, equation 139.

of demand is the ratio of an infinitesimal relative change in quantity to the associated relative change in price, halving the base on which the relative changes in quantity are computed doubles the absolute value of the elasticity of demand. Similar effects are present when the two commodities differ in importance.

The difficulties and limitations of demand studies in an approximately stationary state can be classified into a limited number of types. The condition of equilibrium effectively prevents the statistical determination of economic parameters, but when there are small deviations in quantities it is sometimes possible to obtain useful estimates of price flexibilities and elasticities of demand. The existence of close substitutes necessitates the construction of composite commodities because the high intercorrelation of prices reduces the accuracy of estimates. Composite commodities are also necessary because the number of variables would otherwise be so great as to be unmanageable. There does not seem to be any simple method by which the relation of the demand for a commodity to relevant characteristics of the distribution of income can be determined statistically. It is usually necessary to neglect many variables which have not shown such variation during the period of time covered by the available data.

Neglect of variables which are always unimportant is usually justified. It is sometimes possible to substitute an index number for a group of such commodities if they are closely related. Neglect of a potentially important variable is more serious. The difficulty is not eliminated by including such a variable in the study because estimates of regression coefficients on a given independent variable are not reliable unless the residuals of that variable on the other independent variables show significant variation. For example, the effects of variations of prices of other commodities on the demands for A, B, and C could not be determined accurately because changes in prices of other commodities are income and substitution effects specified to be of little importance. If in the future the quantities of A, B, and C should stabilize and quantities of other commodities should vary, changes in the price of A would depend entirely on price changes which must be neglected in current studies. These difficulties are superimposed upon those caused by changes in population, technological progress, and other important dynamic changes which give rise

to difficulties in realistic studies of demand and supply.

Variations in Both Quantities and Prices

Statistical demand and supply curves can be determined from the same data under certain conditions. This will be illustrated in connection with a model patterned after the potato market of the United States in which all difficulties connected with secular and cyclical changes are abstracted from in order to concentrate attention on other problems. In addition, it is assumed that the data are complete and accurate and that there are no imports or exports. First suppose that the price of potatoes in the United States is determined on the basis of an unchanging demand function in which the quantity of potatoes produced in a given crop year is expressed as a single-valued function of a certain type of average delivered price of potatoes and similar prices of a limited number of related commodities. Similarly, production plans are based on an unchanging supply function in terms of farm prices, but the actual production differs from that which is planned for any year because of variations in the weather. Logarithms of ratios of actual to planned production are specified to be random and normally distributed. Under these conditions the multiple correlation of the quantity of potatoes produced with the appropriate delivered prices is perfect and the parameters of the demand function can be determined exactly from a small number of observations where the forms of both demand and supply functions are linear in terms of logarithms. Similarly, the regression coefficients of the logarithm of potato production on the logarithms of the farm prices of the appropriate commodities are maximum likelihood estimates of the parameters of a specified supply function. When the ratio between average farm prices and the corresponding delivered prices is constant, the use of the same price series for both the demand and the supply studies would change only the constant term, the logarithm of a factor of proportionality. This specification is an oversimplification of the one used by Schultz in his extensive studies of the demand for agricultural products.¹⁵ Schultz points out that it describes approximately the actual responses of potato producers as shown by a study by

¹⁵Schultz, The Theory and Measurement of Demand, pp. 75-80.

Hartkemeier.¹⁶ Schultz also gives a diagram of both the theoretical process and a more realistic process of adjustment following a disturbance to equilibrium under these conditions where there are no further disturbances.¹⁷ When the demand and supply functions are linear in logarithmic form, the condition necessary for return to equilibrium is that the elasticity of the supply curve must be less in absolute value than that of the demand curve.

A better approximation to reality is obtained by applying the foregoing specification except its exclusion of imports and exports to comparatively small subdivisions of the United States which may be called counties for convenience. These smaller markets should be so small that costs of transportation within counties are not important. In this case "international trade" between counties is subject to changes from period to period. In the case of the United States it seems realistic to specify that the correlation between effects of weather on yields by counties is low, especially when counties in which imports or exports would involve important transportation costs are considered. Uncorrelated effects of the weather on actual potato production causes the geographic pattern of production of potatoes to differ from the optimum on the basis of the price structure on which the production plans were based. In addition such shifts in geographic patterns of production make it impossible to describe the relevant characteristics of the distribution of prices by means of an index of farm prices and another of consumer prices for each commodity. Thus a particular type of delivered price of potatoes is not a single-valued function of the national supply and related prices, and the planned production for the coming year is not a single-valued function of index numbers designed to represent the present farm price structure for the whole economy even when similar statements are true for every individual county. Hence, in macrodynamic studies which pertain to an entire country, index numbers with which the economist must represent economic variables need not be perfectly correlated even when the data are complete and accurate and the local prices and production plans are uniquely determined

¹⁶Harry Pelle Hartkemeier, "The Supply Function for Agricultural Commodities: A Study of the Effect of Price and Weather on the Production of Potatoes and Corn," University of Missouri Studies, VII, No. 4 (1932).

¹⁷Schultz, The Theory and Measurement of Demand, p. 79.

on the basis of an extremely simple specification. In the case of the simple model under consideration it would seem more reasonable to specify that observations on the appropriate index numbers of production and prices are random samples from a normal multivariate universe. Even when all relationships are linear in terms of logarithms, no argument based on sampling theory leads to the preference of any of an infinite number of planes as the single-valued relationship on the basis of which to estimate price flexibility and elasticity of demand for potatoes. The choice of such a regression plane, when the data are complete and accurate by hypothesis, would imply definitions rather than estimates of these economic parameters.

If the deviations from a single-valued relationship inherent in the use of index numbers, abstraction from transportation costs, and omission of all but a few important variables are negligible in comparison with technical errors in the available data, the specification of single-valued relationships except for errors might be appropriate for some purposes. This type of reasoning may have been the basis for Schultz's preference for the specification of single-valued relationships except for errors,¹⁸ but his references to pure theory suggest that it would be reasonable to expect the appropriate index numbers based on complete and accurate data to be perfectly correlated. It has been shown that even the mathematical theory of exchange does not lead to that conclusion for the simplest of tenable specifications with which to approximate the conditions of macrodynamic demand studies. Nevertheless, it is probably worth while to summarize the methods by which maximum likelihood estimates of parameters of single-valued relationships can be derived when all variables are subject to random normally distributed errors.

The Weighted Regression

The term weighted regression is here used to refer to the type of regression in which the coefficients are maximum likelihood estimates of the corresponding parameters of a single-valued relationship among a set of ideal series when observations on all of the variables are subject to random normally distributed errors. Schultz called this type of regression the mutual regression.¹⁹

¹⁸Ibid., pp. 146-49. ¹⁹Ibid., p. 149, n. 34.

The set of techniques known as confluence analysis developed by Frisch²⁰ and the restatement and extension of Frisch's arguments in terms of sampling theory by Koopmans²¹ are also related to this type of regression. Errors or disturbances in a set of economic series are discrepancies between available series and ideal variables which presumably would be perfectly correlated. It is extremely difficult to define a set of ideal series in such a way that the definition is useful in approximating for each series the dispersion of the error components as the foregoing examples have shown. Since it is necessary to estimate the relative accuracy of each series in order to determine the coefficients of the weighted regression, serious attention should be given to the problem of definition of errors before the decision to determine the coefficients of a weighted regression is made. One of the most inconsistent uses of the specification of closed sets except for "errors" occurs in the regression analysis of time series. Time is used in many studies as a catch-all factor—an index number designed to represent neglected variables. No sane economist would contend that a smooth function of time is a perfect measure of the effects of neglected variables, but it seems to be the accepted practice to specify that time is free from error.²² It would be more nearly consistent with the undue emphasis on "errors" to specify that all of the errors are in the catch-all variable, but fortunately for economic research, this is never considered seriously even by the most ardent exponent of the device of weighted regressions. Koopmans' insistence on the importance of underestimating rather than overestimating the error variance of a variable which has unimportant effects on the others is another example of the lack of confidence of advocates of single-valued relationships in the implications of the principles of their own system of analysis.²³ In its extreme form this advice implies that a variable

²⁰Ragnar Frisch, Statistical Confluence Analysis by Means of Complete Regression Systems (Oslo: Universitetets Økonomiske Institutt, 1934).

²¹T. Koopmans, Linear Regression Analysis of Economic Time Series (Netherlands Economic Institute, No. 20 [Haarlem: De Erven F. Bohn N. V., 1937]).

²²Schultz, The Theory and Measurement of Demand, pp. 717-18.

²³Koopmans, op. cit., p. 112.

which is of comparatively little importance in the analysis should not be taken as the dependent variable. As a general rule, subject to exceptions when the net relationship of the "unimportant" variable with one or more of the others happens to be of peculiar interest, Koopmans' advice is worthy of consideration even when it is inconsistent with the determination of regression coefficients on the basis of the relative error variances in all of the variables.

Neither Frisch nor Koopmans discusses the relative usefulness in economic research of single-valued relationships except for errors as compared with the elementary regressions. Both writers imply that the main purpose of regression analysis of economic series is to estimate the parameters of single-valued relationships which would exist in the absence of errors. Contrary to opinions which might be formed on the basis of available treatments of weighted regressions, the presence of errors in all of the variables is not sufficient in itself to establish the applicability of the weighted regression. If the purpose of the investigation is to estimate the value of the dependent variable from observations on the independent variables, it can be shown that the elementary regression is superior to the weighted regression even when the weighted regression coefficients are maximum likelihood estimates. The best case for the use of weighted regressions can be made when independent variables are controllable in the sense that something can be done which will make one or more independent variables larger or smaller than they otherwise would have been. A change in tariff rates is a possible example. In general, it would seem that the relevant relationship is that which expresses the series in which changes are to be predicted as a function of the controllable or observable components of related series.

Weighted Regressions for a Simple Specification

Let each of a set of observed series be a linear function of $r + 1$ normally distributed component variables of the form

$$x_i = a_{i1}y_1 + a_{i2}y_2 + \dots + a_{ir}y_r + c_iv_1. \quad (1)$$

The coefficient of each of the y 's differs from zero in at least two of the x 's while each of the v 's is peculiar to only one x and is usually specified to be the error component or disturbance. In

case it is more appropriate to specify that the error components in a set of variables are correlated, one or more terms in the y 's can be added to the v 's and the sums can be specified to be errors which are, of course, correlated. This is a reasonable specification when two or more series have been deflated by the same index number, but in most cases, it merely increases the difficulties which always arise in the process of defining errors. The method of fitting weighted regressions in the case of correlated errors is presented by Koopmans.²⁴ It is usually better to treat correlated errors as the effects of neglected variables in order to place emphasis on whether or not the effects of omitted variables are really unimportant.

When the relationship among a set of variables is not perfect, it is impossible to set limits which will include the value of the regression of the dependent variable on any independent variable when an additional variable is added to the set. For convenience consider any regression in a set of seven variables solved for x_1 . Let x_0 be constructed in such a way that each observation is proportional to the distance in a direction parallel to the x_1 -axis from the corresponding observation point to any arbitrarily selected plane in seven-space. The coefficients of regression of x_1 on the other variables in the new closed set of eight variables are identical with the corresponding coefficients of the arbitrarily selected plane. Hence, any regression coefficient (elementary or weighted) may either increase or decrease without bound upon the addition of a neglected variable to the set. In view of this conclusion, the mathematical treatment of the specification of single-valued relationships can be limited to cases in which the k variables form a closed set. This is accomplished by specifying that the rank of the determinant of the coefficients of the y 's in the system of equations (1) is r and that r is less than k . Geometrically this means that the single-valued relationship among the systematic parts is the intersection of $k-r$ planes in k space. It is also convenient to specify that all variables are measured in units proportional to their known disturbance standard deviations. The use of deviations from sample means simplifies the maximum likelihood solutions. The justification for this choice of origin is the same as for the determination of elementary regression coefficients.²⁵

²⁴Koopmans, op. cit.

²⁵Schultz, The Theory and Measurement of Demand, pp. 734-35.

The Method of Maximum Likelihood and
Single-Valued Relationships

Maximum likelihood estimates often maximize the appropriate frequency density function at the sample point with respect to unknown parameters. For example, in random samples of n observations from a normal universe, deviations from the mean of the universe are considered to be errors of sampling. The frequency density is a decreasing exponential function of the sum of squares of sampling errors. When this function is maximized by minimizing the sum of squares with respect to the unknown mean of the universe, the arithmetic mean of the sample is obtained as the maximum likelihood estimate. Geometrically this is equivalent to minimizing a distance in n dimensions. Consider the sample point in relation to a similar point in which there are no errors. The coordinates of the latter point are all equal. Hence the problem is to choose the point which is closest to the sample point subject to the restriction that all of its coordinates are equal. The locus of such points is a line through the origin with equal direction cosines. The common value of the direction cosines is the reciprocal of the (positive or negative) square root of n because the sum of squares of any set of direction cosines must be unity.

The geometrical interpretation of the problem of estimating the mean of a normal universe is analogous to the problem of estimating the parameters of a single-valued relationship among a set of variables when the systematic parts are all proportional and all variables contain independent normally distributed random error components—the simplest case of multicollinearity. Since variables can always be expressed in units proportional to the known standard deviations of their disturbances, only the case of equal disturbance standard deviations need be discussed. Consider a sample of n observations on k variables represented as a scatter of points in k -space. The single-valued relationship is a line in k -space whose parameters are to be estimated by the method of maximum likelihood. As in the case of the mean, maximum likelihood estimates of the coordinates of an errorless point determine the foot of the perpendicular from the sample point to the line of relationship. Since the parameters of the line are unknown the distance from the sample point to the line must be expressed in terms of the unknown parameters of the line. In order to maximize the frequency density function for the sample as a whole, it is neces-

sary to minimize the sum of squares of these distances with respect to the unknown parameters.

When the single-valued relationship is a plane instead of a line, similar arguments are applicable. When units are chosen proportional to the standard deviations of independent normally distributed error components, the problem of obtaining maximum likelihood estimates of the parameters of the relationship among the systematic parts is reduced to the minimization of the sum of squares of perpendicular deviations from the plane of relationship with respect to the unknown parameters. Geometrical representations of solutions for such problems are lines and planes of best fit in the least squares sense. These depend only on the configuration of the observation points and are therefore independent of the choice of coordinate axes.

Let m represent the square array of variances and covariances among the x 's which is known as the moment matrix. The characteristic equation is derived by subtracting m , an undetermined constant, from each of the diagonal entries of M and setting the determinant of the resulting matrix equal to zero.²⁶ Let M_j represent the characteristic moment matrix when the j th root in order of size is substituted for m . The general solution of the equations required for the minimization process for the case of proportional systematic parts shows that maximum likelihood estimates of the direction cosines of the line of relationship are obtained by normalizing the elements of any row or column of the adjoint of M_k .²⁷ Similarly, normalizing the elements of any row or column of M_1 determines a set of maximum likelihood estimates of the parameters of the normal form of the plane of relationship when the systematic parts form a closed set with no closed subsets. These solutions are determinate only when the appropriate characteristic roots are distinct.

The important properties of the solution can be explained in terms of a new set of axes chosen in such a way that all covariances vanish. Excluding cases in which two or more of the new variables have equal variances this choice of axes is unique. Let z_1 repre-

²⁶L. L. Thurstone, The Vectors of Mind (Chicago: The University of Chicago Press, 1935), pp. 1-43. This is the simplest available treatment of mathematical topics related to the characteristic equation.

²⁷Koopmans, op. cit., p. 66.

sent the new variable whose variance is i th in order of size.

Consider the problem of determining in order the coefficients of the k hyperplanes in normal form such that the variance of the perpendicular deviations from the i th plane is a minimum subject to the condition that the i th plane is perpendicular to planes previously determined. It will be shown that the required equation is $z_1 = 0$. Let h_{1j} represent the i th coefficient in the normal form of the equation of the j th plane. The sum of squares of perpendicular deviations from this plane is

$$f = \sum (h_{1j}z_1 + h_{2j}z_2 + \dots + h_{kj}z_k)^2 - nm(h_{1j}^2 + h_{2j}^2 + \dots + h_{kj}^2 - 1). \quad (2)$$

where the last term, which vanishes identically, is added to impose the condition that the h 's must be direction cosines and nm is a Lagrange multiplier. Setting the partial derivatives with respect to the h 's equal to zero and dividing by n results in a system of k homogeneous linear equations of the form

$$h_{1j}(m'_{11} - m) = 0, \quad (3)$$

where m'_{11} is the variance of z_1 .

The determinant of the system is the determinant of the characteristic moment matrix of the z 's. Since the system of equations is homogeneous, the determinant must vanish in order to avoid unsatisfactory trivial solutions. In this special case it is apparent that the k variances are the roots of the characteristic equation. When the i th variance, m'_{11} , is substituted for m , all of the h 's vanish except the i th to which any arbitrary value may be assigned. The normalizing process reduces the absolute value of this coefficient to unity. Hence, the planes determined on the basis of the minimizing process are the coordinate planes of the system of z 's. Since these planes are independent of choice of axis, the same planes are determined from the same characteristic roots if the minimizing process is carried out in terms of the x 's. It is evident that the planes are mutually perpendicular and that the solution is unique. It can now be shown that the variance of the perpendicular deviations from the i th plane is a minimum, subject to the restriction that the plane must be perpendicular to all planes previously chosen. The condition that two planes are perpendicular to each other is that the sum of products of corresponding coefficients must vanish. Since the first $i-1$ planes in-

volve only the first $i-1$ z 's, any normalized function of the last $(k-i+1)$ z 's determines a plane perpendicular to the first $i-1$ planes. It can now be shown that z_1 has less variance than any other function of this class. The variance of a normalized function of the last $(k-i+1)$ variables is a weighted mean of the corresponding variances. A weighted mean of a set of numbers is minimized by allocating weights of zero to all numbers which exceed the smallest of the set. Hence, the normalized linear function of the last $(k-i+1)$ z 's which has the least variance is z_1 .

When it is appropriate to specify that the systematic parts of a set of variables form a closed set with no closed subsets, the single-valued relationship as determined by the method of maximum likelihood is the plane $z_1 = 0$. Advocates of weighted regressions take for granted that this type of regression should always be used except in cases of multicollinearity. The most extreme case of multicollinearity has been discussed. In intermediate cases, the single-valued relationship among the systematic parts is the intersection of $k-r$ planes in k -space where r is greater than one and less than $k-1$. This relationship, determined by the method of maximum likelihood as before, is the intersection of the $k-r$ planes $z_i = 0$ where i ranges from 1 to $k-r$. None of the coefficients of the functions which express the z 's in terms of the x 's are separately reliable even under conditions favorable to the reliable determination of the intersection. The unreliability of the individual coefficients has led to the common statement that regression coefficients have no meaning in cases of multicollinearity where variables in the set satisfy more than one linear relationship. A common procedure in economic studies is to choose from the multicollinear set of variables that subset which most nearly approaches a closed set with no closed subsets. It is interesting to contrast this method with that of multiple factor analysis in which multicollinearity is a necessary condition for success.²⁸

When it is appropriate to specify that the systematic parts are proportional to each other, z_k is the best estimate of their common component according to the criterion of maximum likelihood. This is probably the most important case of the z 's because the existence of multicollinearity among economic variables is common.

²⁸Thurstone, op. cit., pp. 74-77.

This characteristic is an important factor in the explanation of the common practice of combining highly correlated variables into averages or index numbers, even when the differential behavior of some of the variables in the set is important for other purposes. This also serves to reduce the number of variables, but in many cases unreliability of regression coefficients is sufficient in itself to justify the use of index numbers. Estimates of the separate effects of independent variables in a highly correlated subset are extremely unreliable even when the combined effects of the whole group of variables is great. In such cases, it may be advisable to abandon the attempt to estimate the separate effects of variables in nearly closed subsets and to specify that the systematic parts of variables in the subset are proportional where each series contains a normally distributed disturbance. The function z_k for this subset can then be used as the most reliable linear function to represent the economic variable imperfectly measured by every member of the multicollinear set. The disturbance standard deviations which are assigned for the purpose of computing z_k should be determined primarily by the degree to which each variable seems to represent the desired economic variable.

The original specification describes the x 's as linear functions of r component variables which are common to two or more x 's where r is the rank of the moment matrix of the systematic parts and the single-valued relationship among the systematic parts is the intersection of $k-r$ planes in k -space. The observations on the systematic parts are restricted to a region of r dimensions defined by this intersection. For example, when r is unity, the observations on the systematic parts lie on a line defined in terms of maximum likelihood estimates as the axis of z_k . Since z_k is influenced only by variation parallel to this line, it is superior to any other linear function of the x 's for the purpose of representing the single common factor. When r is two, observations on the systematic parts are restricted to a region of two dimensions. The parameters of this region are not unique since it is determined by any pair of hyperplanes which intersect in it. The variables z_k and z_{k-1} are uncorrelated functions of the x 's influenced only by variation parallel to a particular region of two dimensions. This region differs from the one which contains all of the observations on the systematic parts only because of errors in estimating z_k and z_{k-1} by the method of maximum

likelihood. Neither z_k nor z_{k-1} has any simple interpretation with respect to either of the two common factors. Under favorable conditions when both z_k and z_{k-1} are used as independent variables in regression analysis, their combined effects correspond closely to the combined effects of the common factors. Devices used in multiple factor analysis for the estimation of the disturbance variances and the separate estimation of the various common factors in terms of the x 's are not of interest for present purposes. The foregoing argument which can be extended to apply to larger values of r for any problem in which such extension is appropriate is consistent with the other aims of the multiple factor method.

Numerical Approximation of Parameters of Single-Valued Relationships

When the specified single-valued relationship is a straight line, maximum likelihood estimates of its direction cosines are proportional to the elements of a row of the matrix M_k . When the number of variables is large, the labor of computing these estimates by ordinary processes is considerable. It is necessary to compute all principal minors of all orders to obtain the characteristic equation. The numerical value of the largest characteristic root must then be approximated by Horner's or some similar method. Finally, this root is substituted in the characteristic matrix and the elements of a row of the adjoint are computed and normalized. This direct method can be replaced by a simpler iterative process which is especially useful when the number of variables is large. Except for choice of units, the iterative method is equivalent to Hotelling's method of principal components.²⁹

When the rows of the singular matrix M_k are multiplied by numbers proportional to the corresponding elements of any row or column of its adjoint, the column sums vanish. When the moment matrix M is multiplied by the same vector, the column sums constitute a vector which is m_k times the original vector. This follows from the fact that diagonal entries of the two matrices differ by m_k . When the original vector is only approximately proportional to a row of the adjoint of M_k , the process is iterative. When

²⁹ Harold Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components," Journal of Educational Psychology, XXIV (1933), 417-41, 498-520.

successive vectors are normalized, their elements converge towards maximum likelihood estimates of direction cosines of the line of relationship which corresponds to the axis of Hotelling's first principal component. The rapidity of convergence of the iterative process is increased by squaring the moment matrix several times before multiplying by the first vector.³⁰ The result obtained from 2^t iterations applied to the moment matrix can be secured by a single iteration applied to the matrix obtained by t successive squarings of the moment matrix.

Hotelling's iterative processes are applicable to the estimation of parameters of the line of relationship. They can be adapted for use in estimating the parameters of the weighted regression—a plane of relationship. When a number larger than m_k is subtracted from each of the diagonal entries of the moment matrix, the characteristic roots of the resulting matrix are all negative. Corresponding characteristic roots of the two matrices differ by a constant amount, but their ranking in order of absolute size is reversed. Since the iterative process converges towards a set of direction cosines based on the largest root in absolute size, it is apparent that coefficients of the weighted regression are obtained.

Reliability of Regression Coefficients

The specification used in preceding sections determines a normal k -variate universe of the observed series in terms of the means and variances of their independent component variables. The frequency density of a normal k -variate universe is a decreasing exponential function of a positive definite quadratic form. The coefficients of the quadratic form can be written in terms of any $k(k+3)/2$ mathematically independent parameters of the universe. For example, they are determined by either the means, variances, and covariances or by the means, standard deviations, and correlation coefficients. The specification of such a set of parameters does not always determine the nature of the relationship among the systematic parts. This can be shown by constructing one universe in which the systematic parts are proportional and another in which the systematic parts form a closed set with no closed sub-

³⁰ Hotelling, "Simplified Calculation of Principal Components," Psychometrika, I (March, 1936), 27-35.

sets. These universes can then be shown to have identical parameters.

Let U_1 be a universe in which all systematic parts are proportional and all disturbances are nonvanishing. This is the extreme case of multicollinearity in which the relationship is a line, the intersection of $k-1$ planes in k -space. Let U_2 be another universe which contains the same component variables, but all of the v 's except v_1 are specified to be systematic parts. This implies that the systematic part of x_1 is the first elementary regression function. Since all of the x 's are the same, the parameters of U_1 and U_2 are identical. It is evident that observations on the systematic parts in U_2 are not confined to the intersection of two or more planes in k -space. This follows from the fact that the component variables specified to be nonvanishing disturbances in U_1 are uncorrelated by hypothesis so that no subset of the x 's in U_2 can form a closed set. Since U_1 is the extreme case of multicollinearity, it follows that all cases of multicollinearity with nonvanishing disturbances are equivalent in the same sense to the same simple type of normal k -variate universe. For universes of this type all elementary regressions are well defined and sampling distributions of statistics computed in regression analysis have been extensively studied. It is important to establish this fact because it must be used in appraising the relative merits of different devices used to estimate the reliability of regression coefficients.

It has been shown that any normal k -variate universe which can be expressed as a system of variables with independent nonvanishing disturbances and proportional systematic parts can also be expressed in terms of variables whose systematic parts form a closed set with no closed subsets. Similarly, any normal universe can be expressed in terms of the extreme case of multicollinearity, but it is usually necessary to specify correlated disturbances. Correlated errors can be completely avoided only if it is possible to replace the diagonal entries in the moment matrix by a set of positive constants in such a way that the resulting matrix is of rank one. In this case, the x 's are expressed in terms of the z 's and the terms in z_k are systematic parts. These examples show that the type of relationship among the systematic parts depends on the choice of components which are specified to be errors. This principle has important applications. When the purpose of a

is to derive an estimating equation, it is irrelevant whether or not the available series have systematic parts which are perfectly correlated. The estimating equation should be derived in a manner consistent with its intended use; the methods of estimating its parameters should be the same as would be appropriate if there were no errors in the independent variables. An observed value of the dependent variable should be considered as a random sample from an array of values associated with a given set of observations on the independent variables. The research worker should specify that the mean of any array of values of the dependent variable is a single-valued linear function of the independent variables. When such arrays are normal, maximum likelihood estimates of the required coefficients are the elementary regression coefficients.

Maximum likelihood estimates of any set of mathematically independent parameters of a normal k -variate universe can be obtained from a random sample of n observations where $n-k$, the number of degrees of freedom among the residuals, must be greater than zero—preferably considerably greater. One complete set of parameters not previously mentioned consists of the means, the standard error of estimate, the regression coefficients (excluding the constant term), and the sampling variances and covariances of the regression coefficients. These sampling variances and covariances are parameters of the normal $(k-1)$;variate simultaneous sampling distribution of the elementary regression coefficients where successive samples contain the same values of the independent variables.³¹ The sampling distribution of a particular regression coefficient when sampling is unrestricted is the sum of the separate normal sampling distributions derived on the basis of particular sets of values for the independent variables. This distribution is leptokurtic and it is inconvenient for practical purposes.³² The condition of stratified sampling simplifies the distribution and, at the same time, permits a simple derivation of an exact test of significance for the more general case.

When successive samples are selected from the same arrays,

³¹Schultz, The Theory and Measurement of Demand, pp. 720-26.

³²Karl Pearson (ed.), Tables for Statisticians and Biometricians (London: Biometrika Office, 1934), Part II, pp. cxliv-cliv.

the sampling distribution of any regression coefficient in units of its estimated standard error is the t-distribution where the number of degrees of freedom is $n-k$. This statement, which is relatively easy to establish for any single set of values of the independent variables,³³ is also true for all possible sets when (1) the universe is continuous, (2) the arrays of the dependent variable are normal and homoscedastic, and (3) the regression is linear. This implies that the observed values of the independent variables are not perfectly correlated because the elementary regression is indeterminate in that case. For all determinate cases, the preceding statements follow from the fact that no closed sets occur in sample values of the independent variables unless all samples contain at least one closed set. Absence of closed sets depends on the property of continuity. For any sequence of finite universes which approach a normal k -variate distribution, the limit of the relative frequency of samples containing closed sets among the independent variables is zero. Hence the t-test is an exact test of significance on the basis of a specification which determines a normal k -variate universe. Before modifying this simple specification, it is advisable to appraise confluence analysis on the basis of the same specification.

Confluence Analysis

The set of techniques known as confluence analysis was devised by Frisch for use in regression analysis.³⁴ Its most valuable device is the bunch map—one graph for every pair of variables in every subset. For example, a graph representing relationships between x_1 and x_2 in any subset of g variables is a set of lines radiating out from the origin in the first and fourth quadrants. The slopes of these lines are the coefficients of x_2 where the g elementary regression equations are solved for x_1 and all variables are expressed in standard-deviation units. The number i is placed near the end of the line or beam which is based on the i th elementary regression. The coordinates of the end of the i th beam are the i th elements of the second and first columns of the adjoint of the correlation matrix with appropriate signs.

Multicollinearity among a set of variables contributes towards shortness and wide spread among sets of beams in a bunch

³³Smith, op. cit., pp. 77-86.

³⁴Frisch, op. cit.

map. Hence, two criteria of desirability in bunch analysis are tightness of the bunch and absence of extremely short beams. Since the coordinates of the ends of beams of any bunch are proportional to the corresponding partial correlation coefficients, the necessary and sufficient conditions for a given degree of tightness and absence of short beams is that the absolute values of the highest order partial correlation coefficients must be sufficiently high. Since this is exactly the same condition which leads to high values of Students's t —a single-valued function of a partial correlation coefficient of highest order—the criteria of tightness of the bunch and absence of short beams correspond more closely to the common tests of significance than one might suppose after reading:

In other words, they [tests practically equivalent to t -tests] will be entirely meaningless as tests of the "significance" of the regression coefficients. If we nevertheless use them we actually take a number drawn out of one hat as an expression for the "significance" of some other number drawn out of another hat.³⁵

Although confluence analysis has much in common with the use of Student's t , bunch analysis has several advantages for the purpose of detecting characteristics which are commonly associated with samples from a multicollinear universe. The full bunch map is a graphic presentation of relationships among all possible subsets of variables. This facilitates the choice of a subset of variables on the basis of a given set of criteria. The t -test is usually applied to only those regression coefficients of highest order which involve the dependent variable and the levels of significance used are often too inflexible for the purposes of the problem which Frisch had in mind. On the other hand, the tilling tables on which complete bunch maps are based involve so much additional labor that many who use bunch analysis for the complete set of k variables omit maps for all subsets. In this case, there is an actual saving in labor spent in computation as compared with the use of the t -test. Bunch maps for use in preliminary investigations are easily made. They probably require less effort than is necessary for computation of test criteria, but bunch maps are more expensive to publish. However, the bunch map presents more extensive information than is contained in the results of the usual tests of significance. Of course, no probability integrals are

³⁵Ibid., p. 180.

provided but probability integrals for simple t-tests are single-valued functions of the appropriate slopes and some idea of their magnitude can be gained from the appearance of the bunches.

The most serious disadvantage of bunch analysis is that mental allowance must be made for the number of observations in the sample. Too much reliance is apt to be placed on results obtained from small samples. Waugh's studies of price flexibility reported by Frisch are good examples.³⁶ In these two studies annual observations covering a period of thirteen years were used. Seven variables are included in each study of which two are first differences and one is time as a catch-all variable. Hence there are only twelve observations on seven variables and only five degrees of freedom among the residuals. In regression analysis of serially correlated economic time series, this is a small sample. Even for random samples from a normal k-variate universe, twelve observations on seven variables are hardly enough to make reliable tests possible. It is unfortunate that Frisch made no reference to the extremely small number of observations on which his extensive bunch analyses are based. However, the studies serve a useful purpose in illustrating an important principle of sampling which is apt to be neglected in bunch analysis.

When the specification states that samples are selected at random from a normal k-variate universe, alternative hypotheses differ only in the magnitudes assigned to one or more of the $k(k+3)/2$ parameters of the universe. The common tests of significance are highly useful when a choice is to be made among such alternative hypotheses provided that they differ in such a way that the test to be applied is the most powerful test.³⁷ For example, the t-test is the most powerful test which can be applied to a given regression coefficient when the values of that coefficient specified by alternative hypotheses are all greater or all less than the sample coefficient tested and all other parameters are the same in all alternative hypotheses.

Tests comparable in scope and efficiency with the common tests of significance under favorable conditions are not available

³⁶Ibid., pp. 147-83.

³⁷Jerzy Neyman, Lectures and Conferences on Mathematical Statistics (Washington: Graduate School of the U. S. Department of Agriculture, 1938), pp. 33-48.

for "realistic" specifications for regression analysis of economic time series. Alternative hypotheses for regression analysis of time series should include provision for neglected and potentially important variables as well as the processes which lead to nonsense relationships so often encountered in regression analysis of economic time series. These differences are not in magnitude alone, but also in the structural relationships of the universe which are absent from the simple specifications on which tests of significance or measures of reliability are based. These important characteristics of alternative hypotheses are "specified" out of existence by Koopmans when he states that "a repeated sample consists of a set of values which the variables would have assumed if in these years the systematic components had been the same and the erratic components had been other independent random drawings from the distribution they are supposed to have."³⁸ It is true that the t-test can be modified for use when the systematic parts are constant, but the modification of the specification does not add to the effectiveness of the test when a choice is to be made among alternative hypotheses which are plausible from an economic point of view--the only useful purpose of any test of significance.

In view of the fact that the major part of the variation of a time series is often identified statistically as the systematic part, the consequence of specifying that such components may be serially correlated is of some importance. Specifications which place no restriction of randomness on the systematic parts but offer no solution for problems related to dynamic changes are only superficially realistic. No statistical specification should be chosen in preference to a simpler one on the basis of resemblance to problem conditions for which neither offers any useful solution. Simplicity of specification is important when the same modification of results must be made in any case.

The work of Frisch and Koopmans has called attention to unrealistic specifications on the basis of which misleading interpretations of the common tests of significance are often made. On the other hand, the devices suggested by these writers are sometimes applied when simpler methods are superior. Since no set of techniques is universally applicable, methods of statistical

³⁸Koopmans, op. cit., p. 7.

analysis should be chosen on the basis of important problem conditions.

A Study of Synthetic Production Functions

Regression analysis of economic series is complicated by variables whose effects are difficult to eliminate because there is no simple method of measuring them. This can be illustrated most simply by the construction of a synthetic example similar to those used in the discussion of seasonal variation in Chapter II. Let the production function of a manufacturing establishment be of the form

$$P = bL^kC^j, \quad (4)$$

where P represents the physical product, and L and C represent the factors of production labor and capital, respectively. Knowing the functional form of the production function, the management could determine the parameters by varying the labor and capital applied and determining the linear regression of $\log P$ on $\log L$ and $\log C$. The same type of regression analysis is appropriate when observations on the three variables are available for a large number of establishments governed by the same production function in which the proportions of the factors of production vary from one establishment to another.

Under these conditions it is possible to introduce a new variable, size, by supposing that the establishments are comparable except that some are larger than others. Mathematically, this is accomplished by multiplying each of the variables P , L , and C by the variable S . If the relationship among P , L , and C is single-valued the observations on p , l , and c all lie on a plane where small letters are used to represent logarithms of the corresponding capital letters. The problem to be investigated is to determine the effect of the introduction of S where the regression coefficients of $p + s$ on $l+s$ and $c+s$ are to be compared with k and j . The introduction of the variable S displaces the observation points by a distance s times the square root of three in a direction parallel to a line through the origin whose direction cosines are equal. If the plane which contains the observations on p , l , and c also contains this line, the observations are moved along the plane and the coefficients k and j remain unchanged. The condition that the plane of relationship contains the line whose direction cosines are equal is that the sum of k and j is

equal to unity, that is, the common production function is homogeneous and of the first degree.

It can be shown that the limit of the sum of k and j as the variance of s increases without bound is unity regardless of the magnitude of the sum before the introduction of S . In fact, the sum of the regression coefficients would approach unity even if there were no relationship among P , L , and C . This is apparent on consideration of the fact that the new variables are $p+s$, $l+s$, and $c+s$. Hence, the fact that the sum of k and j is unity in a study involving establishments of widely varying size is not good evidence that the individual production functions are of the first degree. Of course, results consistent with the convergence of the sum towards unity with increasing importance of the size variable are not good evidence against the hypothesis that the individual production functions are of the first degree. The important principle illustrated by this example is that results which agree equally well with two alternative hypotheses cannot be used as evidence to support a choice between the two. The choice must be supported by arguments which are independent of the results in question.

The presence of size as a disturbing variable is typical of statistical studies of economic relationships. The individual observations on the variables used in regression analysis rarely refer to comparable institutional settings and there is often no satisfactory means of reducing the individual observations to comparability. In the foregoing example, the proper adjustment would be to divide by S ; however, such adjustment is usually impossible because S is unknown. Of course, there is usually more than one characteristic with respect to which the individual observations are heterogeneous so that the appropriate remedy is often difficult to formulate.

This study of synthetic production functions has nothing to do with the question as to whether production functions should apply to individual firms of comparable size, to firms of varying size, or to industry totals. It shows, however, that the interpretation of the coefficients k and j for each of these cases is different. For purposes of exposition it is assumed that such functions should be applied to comparable firms of the same size or to observations on the same firm obtained under the same external conditions where the proportions of factors of production

are changed experimentally. With these restrictions it has been shown that introduction of size does not change estimates of k and j if their sum is unity.

The relationship between the factors of production and the physical productivity in the form of industry totals is not a simple one even when the same production function applies to every firm in the industry. The use of industry totals in the equation of the individual productions does not seem to be so objectionable as might appear from consideration of the condition

$$(L_1^k C_1^j) + L_2^k C_2^j = (L_1 + L_2)^k (C_1 + C_2)^j \quad (5)$$

which must be satisfied if the industry consists of two firms. The use of industry totals is equivalent to the use of the arithmetic means of the variables P , L , and C and it is easy to show that the equation of the individual production functions is satisfied by the corresponding geometric means. The statistical derivation of the constants k and j from industry totals when the number of firms vary from industry to industry introduces the same type of difficulty that was encountered for the case of firms of varying size even when the firms are all subject to the same production function.

The Measurement of Marginal Productivity

The study of synthetic production functions in the preceding section serves as a basis for discussion of studies in the measurement of marginal productivity made by Douglas and his associates and summarized in a recent article.³⁹ Some of these studies are based on time series and others are cross section studies. Only those of the latter type are considered in this section. In addition to studies previously published, five studies based on industry totals for the Commonwealth of Australia and two of its subdivisions, Victoria, and New South Wales are reported for the first time. All of these studies have certain characteristics in common which can be summarized briefly:

1. All of the variables are highly correlated so that the scatter of observation points in terms of logarithms is largely

³⁹Paul H. Douglas and Grace T. Gunn, "Further Measurements of Marginal Productivity," The Quarterly Journal of Economics, LIV (May, 1940), 399-428.

confined to a region in the vicinity of a line whose direction cosines are equal.

2. No matter which variable is taken as the dependent variable, the sum of the regression coefficients is approximately unity. This is characteristic of any set of observations obtained under conditions determined by a homogeneous production function of the first degree.

3. The logarithmic variances are approximately equal. This and the preceding points strongly suggested by the bunch map and summary measures presented, have been checked with the authors' work sheets.

4. In all of the studies there is an approximate correspondence between k and the ratio of wages to value added.

The authors' discussion of the data and of special circumstances help to explain some of the cases in which more or less unexpected results appear. The results obtained from industry totals were checked by converting to an average per establishment basis. The latter results were not presented in detail, but the agreement between the regression coefficients with those based on industry totals is remarkably close. It is probably safe to assume that the average size of establishment varies considerably from industry to industry.

The scatter of observation points about the line with equal direction cosines is sufficient in most cases to determine statistically significant elementary regression coefficients. Although some of the studies are based on large numbers of observations, the scatter is relatively small. Hence, it is necessary to advance reasons not connected with the data if the results are to be applied to combinations of labor and capital outside the range of observations. Only a small portion of the regression surface is determined by the data and there is no statistical evidence that the relationship outside this limited range of values is of the form specified.

As suggested in the examples of synthetic production functions in the preceding section, it is not clear whether the results are to be interpreted as applicable to industry totals or whether they are used to approximate the production functions of individual establishments. The authors say:

If we assume (a) that we have the production function for the average firm, (b) that wages are determined by the marginal

productivity, then for the individual firms under pure and perfect competition the coefficients of elasticity shown above give the elasticity of demand for labor. An increase of one per cent in wages would have been expected to cause a decrease of from 2.5 to nearly 4.0 per cent in the number of workers employed, with the coefficient for the later years ranging between 2.5 and 3.0.⁴⁰

The extension of the ordinary interpretation of regression coefficients to the derivation of other economic parameters necessitates a thorough examination of the implications of the disturbing element of size and number of establishments in the industry figures. This is not merely a question of the reliability of regression coefficients (on the assumption that the regression function is known to be of the specified form) as determined either by t-tests or by bunch analysis. It also involves an economic interpretation of the separate variables and of the statistically determined relationship. Since this interpretation involves the identification of the regression function in terms of industry totals with a production function of an individual establishment, it is necessary to consider the effects of size and number of establishments on statistical determination of the form and parameters of the individual production functions in case the sum of k and j is not equal to unity or in case the production functions are not of the form specified. This latter would be difficult to detect because size and number of establishments introduce an appearance of linearity in the logarithmic regression functions.

⁴⁰ Ibid., p. 418.

CHAPTER V

REGRESSION ANALYSIS OF ECONOMIC TIME SERIES

Choice of Components for Use in Regression Analysis

In order to measure statistically the relationship among two or more time series, it is necessary to specify which components of the various series are related, regardless of the fact that such components may not have the characteristics of random series which form the basis of most manageable statistical specifications. All techniques used in the regression analysis of time series should be subordinate to this principle. It is well known that nonsense correlations are often obtained in the analysis of time series, but there have been relatively few investigations designed to throw light on the distinction between a legitimate search for relationships among serially correlated components of related time series and the nonsense correlations described by Yule.¹ Yule's analysis shows that the sampling distribution of the simple correlation coefficient derived from short sections of certain types of uncorrelated time series containing trends and cycles is U-shaped even for moderately large samples. This fact implies that a given measure of correlation derived from serially correlated time series has much less value in support of inferences with respect to causal relationships than does a similar measure derived from random series because high measures of correlation occur with greater relative frequency in the correlation analysis of time series.

This principle of decreasing evidence value with increasing degree of serial correlation is applicable in many problems of economic research. Increasing interest in dynamic economics in recent years has resulted in the formulation of many economic re-

¹G. Udny Yule, "Why Do We Sometimes Get Nonsense Correlations between Time Series? — A Study of Sampling and the Nature of Time Series," Journal of the Royal Statistical Society, LXXXIX (1926), 1-64.

relationships which involve functions of time. Some of these formulations are satisfactory, but others merely use time as a catch-all variable. Dynamicizing economic relationships in this way implies that there are important neglected variables which might affect adversely the reliability of extrapolations. Inclusion of an independent variable whose chief merit is that it has a particular type of trend which permits the investigator to dispense with the catch-all factor does not solve the problem of nonsense correlations. For example, it is difficult to determine the extent to which the animal units series² used by Schultz are trend factors and to what extent they are important for their own sake. It seems probable, however, that they are of service in both capacities. Difficulties pointed out in connection with the animal units series are typical of problems which arise in connection with regression analysis of economic time series. It is necessary in every study to consider the probability that series which are apparently useful may seem so merely because their secular or cyclical movements resemble corresponding components of other variables.

The mere existence of similarity between trends, seasonal components, or cyclical components of two time series has comparatively little value as evidence of a causal relationship. This should not be construed to mean that temporal coincidence of fluctuations in two or more series is always irrelevant when it is connected with any of these components. For example, the high positive correlation between production of eggs and sales of ice cream has little value as evidence of a direct causal relationship between the two variables even though the correlation between the two series is probably higher than the correlation between production of automobiles and production of steel. In the latter case a causal relationship can be demonstrated by logical arguments.

The choice among alternative hypotheses is a problem in inductive inference. For purposes of exposition it is convenient to discuss the solution of this problem in terms of existence and productive probabilities.³ Existence probability refers to the probability that the causal system exists, and should be identified

²Henry Schultz, The Theory and Measurement of Demand (Chicago: The University of Chicago Press, 1938), p. 243.

³Ibid., p. 128.

with a function of the a priori evidence which is available in support of the particular hypothesis. Productive probability refers to the probability that the observed results would be obtained if the hypothesis were true. The latter is the probability which is maximized in connection with the method of maximum likelihood. For problems in which both types of probabilities can be evaluated, it is appropriate to prefer that hypothesis for which the product of the existence and the productive probabilities is a maximum. Thus, if an economist thinks that there are important theoretical grounds for preferring a particular hypothesis, he may be justified in choosing it even though there are other causal systems in which the observed results occur with greater relative frequency. This procedure which is consistent with the theory of inductive inference often leads to difficulties because the weight which the economist gives to considerations based on economic theory has little evidence value to another economist whose interpretation of economic theory differs from that of the investigator.

When two economists disagree on the weight which should be given to principles of economic theory in the choice of alternative interpretations of statistical results, further discussion is limited to analysis of statistical evidence. This implies that the evidence value of statistical correlations is admitted only when the productive probability based on the hypothesis of one economist is greater than the productive probability based on the hypothesis preferred by the other. It is difficult to apply this criterion even though it is identical with the criterion of maximum likelihood when alternative hypotheses differ only in the value of a single parameter. Sampling distributions for types of alternative hypotheses appropriate for regression analysis of time series are not available, but sometimes differences are so great that this does not present great difficulty.

Since statistical correlations have comparatively little evidence value when the empirical relationship involves components which are not random, it is especially important to support the results of regression analysis of time series by logical arguments. This does not mean that the theory of sampling should be abandoned; it merely means that one should avoid overemphasizing the particular hypothesis that the observed series were selected at random from an uncorrelated universe. There are usually other more plausible alternative hypotheses among which a choice must be made. The

evidence should be presented in a manner designed to aid the economist in making a choice among plausible alternative hypotheses. A research worker can always formulate two or more hypotheses which he considers most plausible and concentrate attention on the presentation of evidence designed to show reasons for his choice.

Relationships among economic time series frequently involve components of related series which are not random with respect to time. Since randomness is a feature of familiar specifications for which exact tests based on sampling theory are available, it is necessary to consider these tests as approximations or else adjust the series until only random components remain. The superficial realism of specifications with constant systematic parts for regression analysis of time series was discussed in the preceding chapter. The absence of statistical specifications on which satisfactory tests for time series can be based is not sufficient justification for making a Procrustean bed of the available tests of significance. Removal of serially correlated components of time series before determining statistical relationships should be based on sound logical reasoning. If the economist believes that the relevant relationship involves serially correlated components, consistency requires the use of these components as variables in regression analysis. Douglas's argument in support of his decision to use original series without adjustment for trend is consistent with this principle.

To insist that in any study of causal relationships the trends must first be eliminated is, in effect, to throw the baby out with the bath. From the long-run point of view, the relative slopes of the various trends are indeed more important than the relative deviations from these slopes. Moreover, these various trends are not entirely independent of each other, and one of the tasks of economic science is to discover what the nature of this relationship is. Nor should the economist be deterred from studying it by the cry that he is indulging in nonsense correlations, when it is patent that the interrelations are not nonsensical. We can, therefore, accept the data as they stand as furnishing valuable evidence upon the actual relationship between L, C, and P during the period studied.⁴

The foregoing quotation is taken from the report of the result of Professor Charles W. Cobb's statistical determination of a production function for the manufacturing industries of the

⁴Paul H. Douglas, The Theory of Wages (New York: The Macmillan Co., 1934), p. 141.

United States from annual observations covering the period 1899-1922. The production function was of the form $P = b_1 k C^j$ where the sum of the exponents was arbitrarily specified to be unity. Douglas discussed the procedures in detail so that careful readers might be in possession of the relevant facts. In particular he emphasized that the same production equation does not apply to every country or to every industry in a given country and that production functions are subject to technological changes.⁵ He also suggested multicollinearity when he stated that "if a wider range of observations were available, it might not only be found that different exponents should be given to labor and capital, but that a different type of equation would be necessary to describe the relationship between the three variables."⁶ In the original study the production function was fitted by dividing through by C, taking logarithms, and minimizing in the P direction. Later studies indicate that values of k and j determined as elementary regression coefficients of log P on log L and log C are almost identical with values originally obtained.⁷

The study of marginal productivity based on time series for the United States will be used as an example to illustrate the types of evidence which are relevant when a choice is to be made between two alternative hypotheses or economic interpretations of statistical results. Douglas evidently believed that the statistically derived regressions had some validity as production functions. On the other hand, Menderhausen has stated an alternative hypothesis which ascribes the results to relationships among three logarithmic trend slopes.

The constant k indicates whether output and labor changed in the same proportion or not. It shows whether the volume of output per unit of labor remained constant or varied during the period under observation. In our case, this ratio changed constantly; it increased. Thus the divergence of k from 1 may be taken as an indicator of an evolutionary change in the volume of output per unit of labor. The particular value of k is due to the fact that simultaneously the industry became more "capital intensive" and that the numerical values of the three logarithmic trend slopes were such that the differences stood in the ratio of approximately 4. k may, therefore, be taken as a kind of measure of the

⁵Ibid., p. 203.

⁶Ibid., p. 220.

⁷Paul H. Douglas and Grace Gunn, "Further Measurements of Marginal Productivity," The Quarterly Journal of Economics, LIV (May, 1940), 402.

evolutionary changes occurring in industry during the twenty-four years.⁸

In fairness to the authors concerned two points should be explicitly stated: (1) Douglas and his associates have since shifted emphasis from production functions based on time series to cross section studies;⁹ and (2) Mendershausen may have prefaced his analysis of relevant evidence with question begging for purposes of exposition only; he later offered support for his criticisms in the form of economic interpretations.¹⁰ For these reasons the following discussion is apt to appear hypercritical, but it will be shown that neglect of the principles involved in this example have led to the use of irrelevant evidence in research.

The first part of Mendershausen's article is devoted to a critical analysis of the Cobb-Douglas study of the production function based on United States data for the period 1899-1922. In this analysis Mendershausen discussed the implications of several of Douglas's abstractions placing special emphasis on the distinction between used and idle capital.

Let small letters represent logarithms of the corresponding capital letters. A production function fitted to p , l , and c (with the restriction that the sum of the regression coefficients must be unity) is geometrically equivalent to the plane determined by the line through the means with equal direction cosines and a line in the lp plane—the regression of p on l when the corresponding values of c are subtracted from both variables. Mendershausen pointed out that if Douglas's hypothesis were appropriate, the elementary regression of p on l and c would coincide with the plane fitted with restrictions. Finding that the elementary regression coefficients were practically identical with the corresponding coefficients of the plane fitted with restrictions, Mendershausen then resorted to a shift in specification. He accomplished this by repeated emphatic references to the deviations from the line about which the observation points cluster as dis-

⁸ Horst Mendershausen, "On the Significance of Professor Douglas' Production Function: A Correction," Econometrica, VII (October, 1939), 362.

⁹ Douglas and Gunn, op. cit., pp. 399-428.

¹⁰ Horst Mendershausen, "On the Significance of Professor Douglas' Production Function," Econometrica, VI (April, 1938), 143-53.

turbances. This arbitrary substitution of one specification for another is a form of question begging which often serves to eliminate the question at issue from further discussion.

In the terminology used by Mendershausen, Douglas's hypothesis states that the relatively small deviations from the line in directions parallel to the elementary regression plane are systematic parts. Except for special cases of rare occurrence one can always find maximum likelihood estimates of direction cosines of a line of relationship if he cares to specify such an extreme degree of multicollinearity. Douglas did not claim that observation points were widely scattered over the plane of relationship; in fact, his discussion shows that the scatter is not great. Mendershausen's demonstration that a linear function fitted to the trend values of p , l and c is almost identical with the Cobb-Douglas production¹¹ depends on the fact (admitted by Douglas) that the observed points cluster closely about a line, but it does not explain why the elementary regression of p on l and c coincides so well with the homogeneous first degree function. The choice between the two specifications should depend on the economic interpretation of the deviations from the line, especially those which are important in the determination of the coefficients of the regression functions.

The reason for the choice of this example can now be explained. Mendershausen recognized the importance of the economic interpretation of deviations from the line in directions parallel to the elementary regression plane of p on l and c .¹² He showed clearly that the slopes of the elementary regression are largely determined by the extreme deviations from the line which represent the years 1908, 1921, 1922. Mendershausen attributed the extreme deviations of the first two points to differences between capital used and capital present while he suggested technological change as an explanation of the third. If these economic interpretations are accepted, the conclusion implied by Mendershausen's analysis follows; there is practically no relevant evidence in favor of the hypothesis that the statistical relationship is a production function. Historical research designed to investigate in detail the adequacy of Mendershausen's economic interpretations would add nothing to the value of the example for the purpose of illustrating

¹¹Ibid., p. 153.

¹²Ibid., pp. 150-52.

principles of statistical inference. The importance of discrimination between relevant and irrelevant evidence should be apparent.

Douglas's report of the original study contains three contour diagrams which show clearly that the observation points are clustered closely about a curve so that the curvilinear correlation between L and C is nearly perfect.¹³ The extreme deviations emphasized by Mendershausen are almost as conspicuous on Douglas's charts as they are on the charts presented by Mendershausen.¹⁴

As indicated above, Mendershausen devoted at least as much attention to the implications of his shift in specification as to his economic interpretation. Unfortunately, the excellent quality of his economic analysis seems to have misled Douglas and his associates into accepting his specification shifting as a relevant type of argument. Their use of bunch analysis and the quotation from Koopmans to the effect that the error variance of an unimportant variable should not be overestimated constitute evidence of such acceptance.¹⁵ It is not necessary to consider the other directions of minimization at all in order to establish the usefulness of the elementary regression of p on l and c. On the other hand, an economically meaningful interpretation of the deviations from the line about which the observation points admittedly cluster is necessary in any case. An elementary regression function represents the mean value of the dependent variable as a function of the independent variables even if the three variables have systematic parts which are perfectly correlated. Thus although Mendershausen may have used bunch analysis and specification shifting for purposes of exposition only, the latter device has proved to be not only unnecessary and irrelevant but also misleading.

Statistical Deflation in Terms
of Regression Analysis

A specification often used in statistical demand studies states that the rate of consumption of a particular commodity in

¹³Douglas, op. cit., pp. 218-19.

¹⁴Mendershausen, "Professor Douglas' Production Function," Econometrica, VI (April, 1938), 149.

¹⁵Douglas and Gunn, op. cit.

a given market is a function of the real price of the commodity and other related variables. The derivation of a series of real prices for use in such a study is usually accomplished by dividing the observed prices by an index of the general price level. This process of statistical deflation is useful if relationships of real prices to other economic variables are more nearly invariant than are the corresponding relationships involving observed undeflated prices. Comparison of the effectiveness of the ordinary process of deflation with that of adjustment by means of regression functions can be made most effectively in terms of simple specifications. For example, it is convenient to specify that all relationships are linear in terms of logarithms.

Let y represent the logarithm of the price of sugar in a given market; x , the logarithm of the rate of consumption of sugar in the same market; and p , the logarithm of an index of the general price level. Let the conditions of supply vary in such a way that observation points representing these three variables exhibit sufficient variation in the neighborhood of the plane of regression of y on x and p to permit satisfactory estimates of price flexibility. Either the partial regression coefficient of y on x or the simple regression coefficient of $(y - p)$ on x can be used under favorable conditions. It is possible to compare the accuracy of estimates derived by the two alternative methods on the basis of the sampling variance of each type of estimate where each set of observations is considered to be a random sample selected from homoscedastic arrays of y . It will be desirable also to consider conditions under which such estimates would be biased.

According to the foregoing specification, when $b_{yp.x}$ is equal to unity the stratified sampling variance of $b_{(y-p)x}$, price flexibility as a simple regression coefficient, is $(1 - r_{xp}^2)$ times as great as the sampling variance of $b_{yx.p}$, price flexibility as a partial regression coefficient. This is merely an illustration of the general principle that substitution of known values of parameters in a regression function increases the accuracy of estimation of the remaining parameters. This is sufficient to establish the superiority of the deflation method for cases in which the price of sugar is proportional to the price index which is used as a deflator, i.e., the partial regression coefficient of y on p in the universe is unity.

The proposition that the price of a given commodity is

proportional to a general price index is of questionable applicability where changes over a business cycle are under consideration. The price of the commodity may be either more or less sensitive with respect to cyclical changes than the average price of the commodities on which the index is based. Such an absence of proportionality often imparts a bias into estimates of price flexibility derived by the deflation method. When the price of the deflated commodity is less sensitive to cyclical changes than is the deflating index $b_{yp.x}$ is less than unity while if the opposite is true $b_{yp.x}$ is greater than unity. When the difference in sensitivity to cyclical changes is great, the mean square error in sample values of $b_{(y-p)x}$, which agrees with the sampling variance only when estimates are unbiased, will be greater than the mean square error in sample values of $b_{yx.p}$ and the deflation method loses its superiority.

Results obtained by the deflation method are sometimes misleading when the deflating index and one or more independent variables are affected by the same neglected variables. Schultz's studies of the relationship between the demand for sugar and business conditions in the United States can be used to illustrate this point.¹⁶ In these studies Schultz introduced w , the logarithm of Persons' "Index of Industrial Production and Trade," and t , time as a catch-all variable, in addition to y , x , and p of the foregoing example.¹⁷ Contrary to expectation, he found the partial regression coefficient of w to be negative for each of the periods studied when either $y-p$ or x was used as the dependent variable.

When the parameter to be estimated is the partial regression coefficient of y on w with x , p , and t "held constant," conditions are most favorable for its estimation by the deflation method when the universe value of $b_{yp.xwt}$ is unity. When the value of this parameter is less than unity, estimates of the regression coefficient of y on w derived by the method of deflation are subject to a negative bias because p is positively correlated with w . Such bias arises when the price index is subject to greater relative cyclical fluctuation than is the price of sugar. The

¹⁶Schultz, The Theory and Measurement of Demand, chap. vi.

¹⁷Primes used by Schultz are dropped here for convenience.

writer made a multiple regression study using Schultz's sugar data for the period 1895-1914. As suggested above, the sample partial regression coefficient of y on p is less than unity, but the discrepancy from the value implied by the deflation method is small. Thus, the suggested hypothesis affords only a partial explanation because the partial regression coefficient of y on w is negative for this period even when the price index is included as an independent variable instead of being used as a deflator.

According to the specification appropriate for the use of weighted regressions it seems reasonable to assume that "errors" in economic series with large cyclical components are correlated. This condition introduces difficulties in addition to those inherent in the problem of putting content into the concepts used in estimating weighted regression coefficients.

For the case in which all errors are specified to be independent, the maximum likelihood estimate of the weighted regression coefficient of y on x is larger in absolute size than the corresponding elementary regression coefficient of y on x and smaller in absolute size than the reciprocal of the elementary regression coefficient of x on y . This property, which is equally true for simple or partial regression coefficients, is the basis for Schultz's upper limit for the elasticity of demand¹⁸ and for the "significance factor" advocated by Frisch.¹⁹

¹⁸Schultz, The Theory and Measurement of Demand, p. 197.

¹⁹Ragnar Frisch, Statistical Confluence Analysis by Means of Complete Regression Systems (Oslo: Universitetets Økonomiske Institutt, 1934), p. 109.

CHAPTER VI

DEFLATION AND THE MEASUREMENT OF ECONOMIC VARIABLES

National Income as an Example

It has been emphasized in preceding chapters that economists are forced to use index numbers to represent economic variables. The process of deflation is used in the construction of many of these index numbers. Hence, it is desirable to illustrate some of the problems of statistical deflation in connection with the measurement of some important economic variable. The measurement of national income was chosen for this purpose because it contains a wider variety of problems than does the measurement of most other economic variables. At the same time, the setting of the problem is familiar and the importance of the criterion of consistency of techniques with purpose is relatively easy to establish.

Definition of National Income

National income is made up of goods and services consumed and net changes in national wealth. Since the components of national income are heterogeneous, it is necessary to determine a common denominator to use as a unit of measurement. One which seems to be useful is the monetary unit. Thus the national income is often expressed in terms of money. National income expressed in terms of current money value of goods and services is dependent on the purchasing power of money. For some purposes it is desirable to express national income in such a way that it is as nearly independent of changes in the price level as possible. This purpose is accomplished by converting the components of national income into real income by a process of deflation. The components can then be added to form what is really an index number of production of all economic goods and services. The figure for real national income differs from the familiar indexes of industrial production in two respects; it is expressed in base year dollars

instead of being converted into a percentage of the base period, and it covers a wider range of goods and services.

The importance of the national income total has been overemphasized at the expense of its components. It is necessary to group amounts representing classes of income which are not strictly comparable for all purposes into broader classes, but it is not desirable to carry the process too far. The components of national income should be available on several bases of classification in order that the total which is most appropriate for a given purpose can be constructed. The problems involved in the definition and measurement of national income and wealth are discussed at length by the Conference on Research in National Income and Wealth in a series of studies.¹

Net Changes in National Wealth

Although it is probably consistent with most purposes for which national income estimates are used to avoid as much as possible the problems connected with the estimation of national wealth and its changes, it is neither possible nor desirable to ignore them completely. The changes in wealth which are most often neglected are those which are caused by changing anticipations with respect to the future. (In the case of individual firms, such changes are sometimes recognized following an appraisal where gains are called appreciation and losses are charged to surplus.) In view of the difficulties of measurement, it is not always desirable to attempt estimates of such gains and losses but they are necessary for certain special purposes.

The discounted or present value of future earnings of any asset is sometimes called its economic value. For any asset there are two useful estimates of loss in economic value during any period. The first is the difference between economic value as estimated at the beginning and a similar estimate at the end of the period. The other figure for depreciation is obtained by subtracting economic value at the end of the period from the value which would have been estimated at the beginning of the period on the basis of information available at the end of the period. Probably the latter approximates more closely the loss in economic value

¹National Bureau of Economic Research, Inc., Studies in Income and Wealth, I (New York: H. Wolff, 1937); and II (Baltimore: Waverly Press, 1938).

determined on the basis of perfect knowledge of the future. The difference between the two estimates, which is called windfall gain or loss, is a credit or charge to surplus when it is not considered to be properly assignable to the period. The process of depreciation of assets is not confined to fixed assets used in the production of economic goods; durable consumers goods are subject to losses in value similar to losses sustained by producers goods.

One type of windfall loss is unexpected obsolescence of assets caused by technological displacement. When this type of loss in value is accompanied by reduced prices of goods produced by the displaced assets, deflated national income may not be reduced if the goods affected are included in the deflating index. There are difficulties involved in decisions with respect to deducting losses in value of machines from income. The increase in technological knowledge is usually worth more than the obsolescence in question, but this may not be reflected in market values. In this respect knowledge is like free government services, public roads, etc.; it may be an improvement in national well-being which is not counted in anyone's monetary income.

It is true that patent rights and monopoly control sometimes prevent gains from technological improvements from being passed on to the country as a free service, but many important technological gains have been passed on in this way. It is safe to conclude also that many of these changes have not been reflected in deflating indexes. The reasons for this conclusion are fairly simple. Decreases in price due to technological changes cause shifts in quantities consumed and in relative prices which would be difficult to measure in the simple case where the change affects methods of making the same commodities as before. The complication of new commodities which are not comparable with the old ones necessitates approximate identification of combinations of the old commodities with equivalent combinations of the new ones if effects of technological changes are to be measured. These difficulties are discussed in greater detail in the section on the process of deflation.

Depletion is another means by which the national wealth is decreased. This change is due to the consumption of exhaustible natural resources. It is exceedingly difficult to determine a reliable figure for this item regardless of the definition adopted. The depletion charges which are deductible for income tax purposes

bear little relation to the loss in economic value of the natural resources, and the derivation of reliable estimates of the diminishing economic value of such resources is practically impossible. The difficulties of making independent estimates of depletion are further complicated by current discoveries of new resources which are related to expenditures for discovery and development.

Value changes also occur when there are shifts in demand or supply which seem to add little or nothing to the national wealth. Shifts in values occur when individual industries are subsidized by tariffs or benefit payments, both of which are important in most countries at the present time. Increases in monopoly power may have similar effects. It is not always apparent whether such changes are reflected in the deflating index and these may affect to a considerable extent the reliability of the results.

Important value changes are caused by changes in taxation. These effects of taxation policy make it desirable to consider the goods and services which are furnished free of charge by the government in order that changes in taxation policy will not be reflected in the national income figures. Evaluation of government services leads to many decisions which are highly subjective and not universally satisfactory. Some of the arguments for and against the New Deal depend on decisions of this sort in support of which it is difficult or impossible to submit satisfactory valuations. Changes in taxation policy also involve the rate of discounting expected future earnings. For example, an increase in the permanent annual tax imposed on a piece of land in the amount of a hundred dollars decreases the economic value of the land by two thousand dollars if the rate at which earnings are capitalized is five per cent per annum.

Non-monetary Income

There are a great many important items of income which are not measured in terms of money. When comparisons which are to be made with the total income figure involve periods or localities in which the proportions of such items are nearly the same, non-monetary items are usually excluded from the totals because of the difficulties involved in estimating their money value. One important class of this type of income is income in kind which is important in the case of certain farm incomes. Another important type is the services of members of the family. For example, when

a man marries his housekeeper, the national income is decreased by the amount previously paid to the housekeeper in wages unless adjustment is made for services of members of the family in the process of estimating the national income. Free services furnished by the government are also non-monetary. This class of income is of considerable importance in the form of educational facilities and the use of public roads and parks. These services have increased in importance in comparatively recent years so that it is important to consider them in any temporal comparison which extends over a long period of years.

Business versus Personal Expenses

Expenses incurred for personal pleasure are for consumption purposes while those which arise in the course of business are expenses of production. This principle is recognized for income tax purposes so that allowance for some business expenses is made in income tax returns. However, the distinction is much broader than the allowances which are made for income tax purposes. A great many important classes of expenditure contain a component which is really a business expense. Farmers' automobiles are both durable consumption goods and factors of production for most of them are used to haul produce or for other purely business purposes as well as for pleasure. Similarly commuters' fares, expenses for the sake of appearance, rental differentials paid by workers for living where jobs are available, and many other similar items have a component of business expenditure in them. Like many other items which are difficult to measure, adjustments for these items are neglected in most estimates, but it is necessary to estimate their magnitudes whenever the purpose of the problem requires a comparison of incomes for classes of recipients in which differences with respect to concealed business expenses are important.

The Use of Deflation in the Estimation of National Income

There are many causes for lack of comparability among measurements of different classes of income in the same period in the same country as well as among incomes of different periods and different countries. When temporal comparisons or comparisons between countries are to be made, it is usually desirable to deflate

the income figures by components before combining them into totals. This is particularly true of net change in inventories—a difference between two wealth items—because it is necessary to deflate both the beginning and closing inventories in order to obtain the net change in real terms. The corresponding procedure is rarely followed in the estimation of the net change in fixed capital because of difficulties connected with windfall gains and losses. A more satisfactory estimate of net capital formation can usually be obtained by subtracting an estimate of total depreciation from the net production of capital goods. The choice of an alternative method of estimating net change in capital goods on pragmatic grounds does not affect the validity of the general principle that differences between wealth totals in real terms are differences in deflated totals rather than deflated differences between totals in terms of the same set of prices.

Principles which are applied in dealing with difficulties which are superimposed in practice can be more easily explained when each type of difficulty is isolated for purposes of analysis. An important example of such a difficulty is the effect of a change in technology which results in a cheaper method of producing a given commodity. The result of such a change is a shift in relative prices due to the increased supply of the commodity affected. The equipment formerly used to make the commodity decreases in value by varying proportions because some of it can be shifted to other uses while completely specialized machinery becomes obsolete and has only scrap value. Materials used to make machinery needed for the new process may increase in value.

Consider a number of individuals with given preference systems or indifference functions where the price of one commodity suddenly becomes cheaper and the change is expected to be permanent. Each individual selects from the infinite number of combinations equal in value to his income the particular one which lies on the highest indifference function. Because such choices affect prices of substitutes of the cheapened commodity, each choice is assumed to be based on the system of prices at a new equilibrium. It seems likely that almost all prices would fall under these conditions, but it is difficult to formulate a unique criterion for deflating total income for the purpose of measuring the increase in real income except for one case. All money incomes are reduced in such a way that every individual after the technological change

is able to buy a set of commodities which he considers to be equivalent to that which he chose to buy before the change. In this case it seems reasonable to specify that no change in real income has occurred. On the basis of this criterion the direction of bias in two simple types of price index numbers can be determined. Laspeyres' index number, which depends on base year quantities, has an upward bias under these conditions because it implies that the price level is unchanged if the set of commodities consumed in the base year can be bought for the same amount in the given year. No individual is able to buy the combination purchased in the base year, but all can buy equivalent combinations. An index number of real incomes based on Laspeyres' index of prices is the value ratio of commodities consumed in the given year to commodities consumed in the base year where both combinations are valued at given year prices—a quantity which is less than unity since the numerator has been shown to be less than the denominator. The expression for the real income index number is obtained by dividing the undeflated income ratio—the ratio of values of combinations consumed in the two periods at prices prevailing in the respective periods—by Laspeyres' index number. In a similar manner it can be shown that Paasche's index number, in which prices are weighted by given year quantities, has a downward bias because it implies that the price level is unchanged if the set of commodities consumed in the given year could have been purchased for the same amount in the base year.

It is easy to overemphasize the importance of the conclusions that the use of Laspeyres' index number of prices leads to an underestimate of real income because of the upward bias of the index number and that the downward bias of Paasche's index number leads to overestimation of real income. These conclusions apply only to index numbers based on complete and accurate data while there are good reasons for believing that commodities which are usually omitted in practice are the ones which are largely responsible for the simple conclusion. In addition, changes in relative prices are not always caused by the introduction of devices which make certain commodities cheaper to produce. Changes in tastes are probably of at least equal importance, and it is doubtful if such changes lead to the same conclusions with respect to bias in simple index numbers. In any case bias in the case of changing tastes depends on arbitrary choice of criteria as well as

abstraction from other problems in order to obtain a simple model.

Obsolescence

Before proceeding to a more detailed discussion of commodities which must ordinarily be omitted in the construction of index numbers, it is worth while to consider the problems connected with machinery which becomes obsolete upon the introduction of new methods of production. There has been a great deal of discussion of the question as to whether obsolescence should be treated in the same way as depreciation which has been anticipated. This general problem can be elucidated in connection with an extreme case.

Consider a plant equipped with costly and completely specialized machinery designed to manufacture a certain product. A discovery that a widely distributed natural resource can be transformed by a simple process into a substitute for the hitherto rare product makes the specialized machinery completely obsolete. Neglecting scrap value, cost of removal, etc., it is clear what the facts are in this extreme case; the problem is to determine the most useful method of reporting them in the national income figures.

The economic value of the machinery just before the discovery was practically zero even though the owner valued it at a million dollars. Some of the difficulties of interpretation of this type of reasoning are due to failure to make a distinction between the capacity to satisfy wants and economic value which is a function of scarcity. Assuming no change in tastes the machinery has the same capacity to satisfy wants as before, but alternative means with which the particular wants can be satisfied are no longer scarce. Therefore, neither the machinery nor the knowledge and resources which replaced it have any economic value, although they make a greater contribution to the national well-being than the machinery alone did before. They satisfy not only the more urgent wants which were satisfied by the products of the machinery but also the less urgent wants which could not be satisfied economically because of the scarcity of the products.

The proper method of obsolescence accounting in the extreme case is not determined by the reasoning used in the example; it depends on what the accounts are intended to show. Suppose that the owner of the machine could foresee the forthcoming discovery

perfectly. It would have been advantageous for him to choose a less specialized type of machine or one which would have been more nearly worn out at the time of the discovery if that were possible. Abstracting from this possibility, the owner should have considered the entire original cost of the machinery as a cost to be charged to the period preceding the discovery. The question as to whether this cost is obsolescence or depreciation is one of accounting terminology. Considering the country as a whole, the loss in economic value may be more than offset by a gain in the standard of living made possible by the discovery. This in turn may be nullified by failure to obtain the advantages of specialized machinery because of risks of similar losses in economic value occasioned by possible future discoveries.

Assuming that the proper compromise between Laspeyres' and Paasche's index numbers can be determined satisfactorily, it seems clear that the unforeseen loss through obsolescence is not adequately dealt with in the deflation process if it is considered desirable to show either a decrease of the same amount in the national wealth or the greater increase in national well being due to the discovery. In the writer's opinion, large unforeseen losses in economic value should generally be charged to surplus because they are not properly chargeable to the period in which they were discovered, and they are not sufficiently regular to be of the nature of an insurance premium. The gains in national well-being which occur through the substitution of plenty for scarcity cannot be measured satisfactorily by ordinary methods in which only scarcity values are used. When national income figures are used as indexes of welfare, allowances should be made for both improvements and losses due to technological changes.

Omission of New and Obsolete Goods and Services

Commodities of increasing importance are sometimes substituted for commodities of decreasing importance by substituting an amount of the new commodity equal in value to the component of the index number made up of the old commodity or by some similar procedure. A more common procedure in the construction of index numbers for use as deflating indexes is to include only those important commodities which are common to both periods—or countries in case international comparisons are to be made. This highest common factor method has the effect of omitting those commodities

which have become obsolete and the new commodities which have originated in the interval between the periods being compared. While it does not follow that the new commodities correspond in any way to the free good in the foregoing example, exclusion of new commodities would exclude innovations which increase the potential standard of living if any had in fact been made during the interval between the periods whose national incomes are to be compared. Arbitrarily setting particular combinations of new commodities equivalent to other combinations of obsolete commodities makes possible the inclusion of both types of commodities in the series of index numbers, but the subjectivity of the process should be recognized. If national incomes are to be used as indexes of welfare, a choice should be made between making allowance for changes in potential standards of living in the construction of index numbers and making allowances for the failure of index numbers to measure the effects of such changes. The writer prefers the latter alternative in the case of important changes; however, the former has advantages for those who insist on making all comparisons on the basis of measures which they consider objective and "scientific." Of course, almost all commodities are subject to gradual change and the distinction is not so sharp as the foregoing discussion implies.

A component of national income which gives rise to many arguments is that of government services. This component has many elements in common with the example used in the study of obsolescence. Free government services have no market value in any familiar sense and yet they satisfy human wants. In a certain sense they contribute to a rise in the potential standard of living in the same way as did the free good in the obsolescence example. There is an important difference, however; government services are paid for by taxes and there is a presumption that some of them are worth their cost to the community. Thus, there is some possibility of obtaining estimates of the economic value of certain government services. There are special difficulties of estimation; some of the services are business expenses as opposed to personal expenses while others are merely transfer payments in which money taken from one person is given to another. In the evaluation of government services, a distinction is sometimes made between (1) services or yields of nationally owned capital or resources which are considered to be economic services and (2) those services

which, like the free good in the obsolescence example increase the national well-being without having economic value. For example, the collective well-being of Chicagoans depends to some extent on the enjoyment they derive from Lake Michigan and from the city parks. Satisfactions derived from the lake might be taken for granted while those derived from municipally improved parks might be considered as the yield of an investment for which estimates of economic value should be determined. For certain special purposes it may be advisable to assign economic valuations to as many government services as possible, but emphasis should ordinarily be placed on those services which are not common to the periods or localities which are being compared.

The Components of National Income

The national income for any country for a given period can be divided into components on the basis of many different classifications. There are three important criteria for deciding which basis of classification is apt to be most effective for a given purpose. The first criterion states that there should be correspondence between the quantity approximated by one component or the sum of two or more components and the economic variable which would be ideal for the investigator's purpose. Another criterion requires that the data from which the magnitude of the various components must be estimated should be reasonably accurate. Some bases of classification might be ideal for purposes of analysis and yet be unsatisfactory because the available data furnish no simple and reasonably accurate method for estimating some of the important components. When real income is to be estimated, it is sometimes necessary to use a third criterion; the basis of classification should be chosen in such a way that satisfactory deflating indexes can be constructed for each of the important components. Of course, it is possible that the three criteria suggested may not be consistent with each other in the sense that each would lead to the choice of a different basis of classification, in which case the choice among bases of classification should depend on the relative importance of the three considerations for the purposes of the problem.

Comparisons of Incomes of Classes of Individuals

In recent years considerable interest has been directed toward the problem of income parity for agriculture. This problem

furnishes good illustrations of most of the kinds of difficulties which are encountered in the measurement of national income. The problem has an item of additional interest in that it is defined by the Soil Conservation and Domestic Allotment Act of 1936. The purpose of the act is declared to be the

. . . . reestablishment, at as rapid a rate as the Secretary of Agriculture determines to be practicable and in the general public interest, of the ratio between the purchasing power of the net income per person on farms and that of the income per person not on farms that prevailed during the five-year period August, 1909 - July, 1914, inclusive, as determined from statistics available in the United States Department of Agriculture, and the maintenance of such ratio.²

The problem of income parity for agriculture is one in which components of national income figures are interpreted as indexes of welfare of economic classes. The passage quoted from the act of Congress is superficially precise in its terminology, but the provisions are actually ambiguous to a pronounced degree. After citing the quoted passage, Stine goes on to say that "we are interpreting this text to require that estimates of current income per person be divided by appropriate index numbers of prices of goods and services constructed on the prescribed pre-War base, and that the results for the current year be compared with the per capita incomes in the base period."³

The provision that income parity shall apply to "income per person on farms" leads to some difficulties if the proportion of non-farm income in the total income of persons on farms has changed. It may be difficult to determine the extent of such change, but the distinction between those who derive their income from farms and those who live on farms is not important for purposes of income parity unless the proportion of non-farm income has changed since the pre-War period. Stine does not emphasize the importance of determining whether the proportion of non-farm income has shifted through the years, nor does he indicate what factors might operate to change this proportion. It is interesting to consider how some of these factors influence the Federal govern-

²Soil Conservation and Domestic Allotment Act, an amendment to the Soil Conservation Act, Public No. 461, 74th Cong., 2d Sess., Sec. 7 (a), (5), cited by O. C. Stine, "Income Parity for Agriculture," Income and Wealth, I, 327.

³O. C. Stine, "Income Parity for Agriculture," Income and Wealth, I, 328.

ment in its efforts to increase incomes of persons on farms. If there has been a net increase in the proportion of investments in land by the non-farm population (for example, by increases in mortgages and owned land held by insurance companies or by increases in the proportion of retired farmers who move to town) this factor would operate to increase efforts to help the farmers. It is probable that many increases in free government services have operated to produce greater improvement in the standards of living of the farm population than in the standard of living for the city population. Among these are improved roads and educational facilities. To the extent that these are taken into account, efforts to help the farmers will be lessened. Unless a great deal of attention is given to measurement of income in kind, decrease in self-sufficiency on farms will be penalized by the provisions of the act. There are many other factors which have affected the proportion of non-farm income received by persons on farms; those considered are meant to be suggestive. It is interesting to note the large element of ethical considerations. The results are affected to an important degree by non-economic elements in the sense that many important items have no market value. The real meaning of the act is apparent only after a decision has been made as to which services and/or satisfactions are to be called income and which are to be excluded—a decision which has important ethical implications when the periods under comparison are characterized by so many obsolete goods and services as well as innovations and increases in government services.

It is not surprising that the question of income parity of agriculture should lead to such ambiguous legislation because it involves arbitrary decisions as to the fair relative standards of living for large classes of the voting population. Public debate of the components of this question on their merits would probably not be politically feasible from the point of view of a member of Congress. The provision that the Department of Agriculture should provide objective criteria of income parity policy in the form of index numbers is a delegation of the legislative function of making the important ethical decisions by defining the main components of income for purposes of the act. Index numbers which are only superficially objective incorporate implicit decisions on each of the debatable points where the maintenance of the appearance of scientific objectivity prevents explicit considera-

tion of each problem on its merits.

The writer is opposed to setting arbitrary relative standards of living for classes of the population because efforts to increase the mobility of economic resources should accomplish the desired purpose with a better approximation to justice for all concerned. Whether this opinion is correct or not, the fact remains that index numbers designed as guides for income parity policy must be based on arbitrary decisions which have far-reaching ethical implications.

CHAPTER VII

CONCLUSION

Statistical Specification and the Scientific Method

Further progress in methodology in economic research depends largely on a better comprehension of the role of specification. When the specification is unrealistic in some important respect, it is unwise to extend the accuracy of estimates of parameters and measures of reliability to many significant figures; it is better to direct the same amount of effort to critical analysis of available specifications. By the latter type of approach it is often possible to determine the direction and approximate extent of revisions which should be made on the basis of important inadequacies. For this reason specifications used in this study have been presented explicitly as oversimplified statements of real problems on which attention should be centered. The degree of realism of a specification is more important than its tractability.

The process of formulating research problems in statistical terms has been neglected in discussions of quantitative methods of research. In fact, most treatments of the scientific method are applicable to the case in which there is universal agreement with respect to specification. The formulation of a good specification depends on the judgment and ingenuity of the research worker and his knowledge of the general field of which his project is a part. If he fails to make the most useful specification because it differs from those which are currently popular he may, through excessive conservatism, lose the opportunity to make a significant contribution to his field. On the other hand, if he makes his specification without reference to specifications of previous studies, his results may be less valuable supplements to the set of explanations of related phenomena which are known as scientific principles.

Specifications for Regression Analysis
and Pure Economic Theory

In its extreme form, pure economic theory is a branch of mathematics whose distinguishing characteristic is its terminology. In many respects the mathematical method is superior to any other for the purpose of investigating the implications of hypotheses as to the nature of relationships among economic phenomena, but its exactness often leads to misinterpretations of the unique conclusions to which it leads. If the results of pure theory are to be useful in the solution of real economic problems it is necessary that the theoretical propositions shall be distinguishable from those of pure mathematics on some basis other than the names which are assigned to the functions used in the economic models; the variables must bear some useful resemblance to observable or controllable elements in the economic system. The foregoing criticisms are not limited to economic propositions derived in terms of mathematical symbols; they apply equally well to verbal arguments based on logical consistency as the sole criterion of validity. In particular, the specification of single-valued relationships must be justified on some basis other than the usefulness of such relationships in pure theory.

The specification of single-valued relationships is often responsible for neglect of important characteristics of economic problems. Even if the index numbers which the economist must use in macrodynamic studies were based on complete and accurate data, it would not be reasonable to specify that they would satisfy a single-valued relationship. For this reason the specification of weighted regressions in the analysis of economic series is equivalent to operational definition of economic parameters instead of a step in the problem of estimating parameters which are already defined. Such specifications should rarely be used because it is usually impossible to give even an approximate definition of the concepts such as "errors" which are inherent in the specification—even in terms of complete and accurate data.

The specification for a regression analysis, which is a hypothesis concerning the nature of the causal system under which the observed data were obtained, should be made on the basis of the investigator's knowledge of the field. Information furnished by the sample should be used with full recognition of the principle that the evidence value of characteristics suggested by the

sample is less than if they were predicted a priori. This is especially true when the series are serially correlated.

Adjustment for Seasonal Variation

Adjustment for seasonal variation is the construction of a function of a time series which is not characterized by variation associated with the seasons. It should be considered as the first step in the regression analysis of related series. Its purpose is to remove the effects of the weather and customs associated with the seasons in order to facilitate comparison of the adjusted series with related series which may or may not be adjusted for seasonal variation.

All of the common methods used to adjust time series for seasonal variation are built up of several component techniques each of which may be combined with components of other methods. New methods can be constructed in this way so as to be especially appropriate for a specification with respect to relationships among components of a particular time series. Thus a moving average can be used to eliminate the effects of the cycle, the arithmetic mean can be used for averaging normally distributed components, the median or average of central items can be used for averaging when extreme deviations from the average are apt to occur, and other variations of methods can be adopted to meet the requirements of the simple specification which is thought to be most realistic. When the problem of seasonal adjustment is approached in this way the flexibility of graphic methods of seasonal adjustment is seen to be a theoretically important property.

Trend Analysis as a Deflation Technique

Adjustment of a time series for trend is the construction of a function of the series which is free from long run movements. Its purpose is to derive a series whose relationships with other series are more nearly invariant than would be the case if unadjusted series were used. Sometimes it is possible to secure equivalent results by including time, a catch-all factor, as one of the independent variables in a multiple regression study. In this case time is used as an index number to represent the effects of neglected variables.

In connection with either of the foregoing methods of dealing with trend, arguments based on the maximum likelihood

property of trend coefficients derived by least squares should be given little or no consideration especially if cyclical components are important. Similarly, computation of standard errors of either trend coefficients or trend functions adds little or nothing to the investigator's background of information on which he must base his interpretations. In the analysis of such series the specification on which standard errors are based and on which the maximum likelihood property depends is so unrealistic in its most important characteristics as to be practically useless. The reason for the rejection of arguments based on sampling theory depends on the principle that sampling theory is only useful as an aid in choosing between alternative hypotheses. Alternative hypotheses in time series analysis differ in the inclusion of different variables, choice of different functional forms, use of different component series in the construction of index numbers, and in other ways none of which corresponds to the variation in the value of a single parameter for which the common tests of significance are most appropriate.

Reliability of Regression Coefficients

When regression coefficients are derived from serially correlated time series, the limited usefulness of measures of reliability is usually recognized. It is almost equally important to recognize the limitations of standard errors when serial correlation is absent. A common characteristic of economic studies is the presence of effects of variables for which satisfactory measurements are not available. In other words, the data are not homogeneous with respect to variables which form the subject matter of the investigation. Under these conditions it should be remembered that familiar specifications state that the data constitute a random sample from a k -variate universe in which the mathematical forms of all relationships are known a priori and that there are no neglected variables. The unrealistic nature of this specification has led to a search for realistic specifications. The most common characteristic of such "realistic" specifications is the avoidance of important economic problems by "specifying" them out of existence. For example, the specification on which Koopmans based his extensive mathematical treatment was shown to be only superficially realistic because the systematic parts of all variables were specified to be constant from sample to sample.

Confidence intervals for regression coefficients are based on a particular causal system. It should be recognized that other causal systems which would lead to much wider ranges of chance variations are not necessarily less realistic than those which statisticians have found to be manageable. This approach leads to a more intelligent interpretation of confidence intervals than does the alternative of trying to translate important conditions not included in the specification into terms familiar to followers of leading mathematical statisticians. Attempts to determine the number of degrees of freedom among n observations on serially correlated time series furnish abundant examples of this type of translation. The student of reliability of regression coefficients should determine as well as he can what important causes might have operated to make his result unreliable. All of these which are not included in his specification should be given separate attention in any discussion of reliability. This procedure is not difficult, but it does involve the abandonment of the search for unique measures of reliability based on standard errors if there are important sources of error besides random errors of sampling or of measurement.

Definition of Economic Variables

Even those who would like to see more extensive use of quantitative methods of research in economics recognize the possible misuse of objective methods. For example, phenomena which can be measured with a fair degree of accuracy may be studied without consideration of any elements which are difficult to measure. This is extremely common in economic research where the elements of size of firm, potential market, and other similar factors are often more important than the variables which can be measured satisfactorily. It was partly for this reason that national income was chosen as an example of an economic variable to be defined.

National income is an index number of production. It is usually expressed in base-year dollars. It is much broader than familiar specialized indexes of production, but no estimate of national income is intended to represent all goods and services. The best criterion for the estimation of national income is conformity with the purpose of the investigation in which the estimate is to be used. For this reason national income figures should

be made available by components on several bases of classification. A research worker can then combine the components in whatever way is best suited to his particular purpose omitting components which are irrelevant and applying special purpose adjustments wherever necessary. Since the components of national income are usually deflated separately, special purpose adjustments are not at all unusual.

In connection with the problem of defining economic variables, the nature of general economic models used in "pure" theory should be considered. Such models are constructed by describing imaginary economies in which economic phenomena occur in conformity with functional relationships. For the purpose of increasing the effectiveness of such models, the economist usually places restrictions on the functions by assuming continuity and specifying the signs of certain first and second derivatives. General economic models which deal with all economic phenomena of an economy lead to conclusions which are really theorems, that is, the conclusions are implied by logical premises which may or may not be realistic descriptions of economic relationships in the real world.

There seems to be a widespread opinion among economists that variables used in empirical studies should be the best feasible approximations to the variables of the same names in a general economic model of scope comparable to that of the stationary state. To the writer this seems inconsistent with the underlying purpose of economic models. General economic models are of service in showing that mutual determination is not inconsistent with sound reasoning, and they are useful for suggesting interesting conclusions. As a basis for empirical studies, however, such general models are too broad in scope and should be replaced by statistical specifications involving definite functional forms of regression functions. Since it is usually necessary to abstract from the effects of remotely related variables and to combine closely related variables into index numbers, it is not realistic to specify single-valued relationships, but the average change in the dependent variable associated with any given system of changes in the set of independent variables can be estimated by means of the ordinary elementary regressions.

Many empirical studies place emphasis on the determination of a single net relationship between two variables. In such cases

other variables are included either because they are important factors in multivariate relationships involving the two variables of greatest interest or they are used to represent important variables with which they were highly correlated during the period covered by the data. Although really important variables are more desirable, the use of variables correlated with them is sometimes the most effective available device for the purpose of facilitating the approximation of the desired relationship. In this connection statistical deflation is the process of applying the adjustments which are appropriate on the basis of the most realistic specification which the research worker can formulate.

BIBLIOGRAPHY

- Bard, William S. "The Use of Statistical Technique in the Administrative Control of Steam Consumption," The Journal of Business of the University of Chicago, X (July, 1937), 251-67.
- Clark, John Maurice. Studies in the Economics of Overhead Costs. Chicago: The University of Chicago Press, 1923.
- Croxtan, Frederick E. and Cowden, Dudley J. Applied General Statistics. New York: Prentice-Hall Inc., 1939.
- Davis, Harold T. "Significance of Analysis of Variance of Time Series," Abstracts of Papers Presented at the Research Conference on Economics and Statistics Held by the Cowles Commission for Research in Economics. Colorado Springs, Colo: Colorado College, 1936. Pp. 99-102.
- Dean, Joel. "Department Store Cost Functions." An essay to be published in a memorial volume for the late Professor Henry Schultz.
- Douglas, Paul H. The Theory of Wages. New York: The Macmillan Co., 1934.
- Douglas, Paul H., and Gunn, Grace. "Further Measurements of Marginal Productivity," The Quarterly Journal of Economics, LIV (May, 1940), 399-428.
- Ezekiel, Mordecai. Methods of Correlation Analysis. New York: John Wiley and Sons, 1930.
- Falkner, Helen D. "The Measurement of Seasonal Variation," Journal of the American Statistical Association, XIX (June, 1924), 167-79.
- Frisch, Ragnar. Statistical Confluence Analysis by Means of Complete Regression Systems. Oslo: Universitetets Økonomiske Institutt, 1934.
- Hall, Lincoln W. "Seasonal Variation as a Relative of Trend," Journal of the American Statistical Association, XIX (June, 1924), 156-66.
- Hart, William L. "The Method of Monthly Means for Determination of a Seasonal Variation," Journal of the American Statistical Association, XVIII (September, 1922), 341-49.
- Hartkemeier, Harry Pelle. "The Supply Function for Agricultural Commodities: A Study of the Effect of Price and Weather on the Production of Potatoes and Corn," University of Missouri Studies, VII, No. 4 (1932).

- Hicks, John Richard. Value and Capital. Oxford: The Clarendon Press, 1938.
- Hotelling, Harold. "Analysis of a Complex of Statistical Variables into Principal Components," Journal of Educational Psychology, XXIV (1933), 417-41, 498-520.
- _____. "Differential Equations Subject to Error, and Population Estimates," Journal of the American Statistical Association, XXII (September, 1927), 283-314.
- _____. "Simplified Calculation of Principal Components," Psychometrika, I (March, 1936), 27-35.
- Joy, Aryness, and Thomas, Woodlief. "Adjustments for the Influence of Easter in Department Stores Sales," Journal of the American Statistical Association, XXII (December, 1927), 493-96.
- Koopmans, Tjalling. Linear Regression Analysis of Economic Time Series. Netherlands Economic Institute, No. 20. Haarlem: De Erven F. Bohn N. V., 1937.
- Kuznets, Simon. "Seasonal Pattern and Seasonal Amplitude: Measurement of Their Short-Time Variations," Journal of the American Statistical Association, XXVII (March, 1932), 9-20.
- Macauley, Frederick R. The Smoothing of Time Series. New York: National Bureau of Economic Research, 1931.
- Marshall, Alfred. Principles of Economics. 8th ed. London: Macmillan Co., 1930.
- Mendershausen, Horst. "Methods of Computing and Eliminating Changing Seasonal Fluctuations," Econometrica, V (July, 1937), 234-62.
- _____. "On the Significance of Professor Douglas' Production Function," Econometrica, VI (April, 1938), 143-53.
- _____. "On the Significance of Professor Douglas' Production Function: A Correction," Econometrica, VII (October, 1939), 362.
- Mills, Frederick C. Statistical Methods. Rev. ed. New York: Henry Holt and Co., 1938.
- Moore, Henry Ludwell. Synthetic Economics. New York: The Macmillan Co., 1929.
- National Bureau of Economic Research, Inc. Studies in Income and Wealth. Vol. I. New York: H. Wolff, 1937. Vol. II. Baltimore: Waverly Press, 1938.
- Neyman, Jerzy. Lectures and Conferences on Mathematical Statistics. Washington, D.C.: Graduate School of the U. S. Department of Agriculture, 1938.

- Osborne, Reba L. "Index of Grocery Chain Store Sales," Survey of Current Business, May, 1937, pp. 15-16.
- Pearson, Karl. "On the Systematic Fitting of Curves to Observations and Measurements," Biometrika, I (1901-1902), 265-303.
- _____. (ed.). Tables for Statisticians and Biometricians. Part II. London: Biometrika Office, University College, 1934.
- Persons, Warren M. "Indices of Business Conditions," The Review of Economic Statistics, Preliminary Volume I (1919), 5-107.
- Pigou, A. C. The Economics of Welfare. London: Macmillan and Co., 1920.
- Preinreich, Gabriel A. D. "Annual Survey of Economic Theory: The Theory of Depreciation," Econometrica, VI (July, 1938), 219-41.
- Rietz, Henry L. (ed.). Handbook of Mathematical Statistics. Chicago: Houghton Mifflin Co., 1924.
- Robb, Richard A. "The Variate Difference Method of Seasonal Variation," Journal of the American Statistical Association, XXIV (September, 1929), 250-57.
- Sasuly, Max. Trend Analysis of Statistics: Theory and Technique. Washington, D.C.: The Brookings Institution, 1934.
- Schultz, Henry. "An Extension of the Method of Moments," Journal of the American Statistical Association, XX (June, 1925), 242-45.
- _____. The Theory and Measurement of Demand. Chicago: The University of Chicago Press, 1938.
- Smith, John H. Tests of Significance: What They Mean and How To Use Them. Chicago: The University of Chicago Press, 1939.
- Spurr, William A. "A Graphic Method of Measuring Seasonal Variation," Journal of the American Statistical Association, XXXII (June, 1937), 281-89.
- Thurstone, L. L. The Vectors of Mind. Chicago: The University of Chicago Press, 1935.
- Tintner, Gerhard. The Variate Difference Method. Bloomington, Indiana: The Principia Press, 1939.
- von Szelski, Victor S. "Comment on 'The Variate Difference Method of Seasonal Variation,'" Journal of the American Statistical Association, XXV (March, 1930), 83-84.
- Yule, G. Udny. "Why Do We Sometimes Get Nonsense Correlations between Time Series? — A Study in Sampling and the Nature of Time Series," Journal of the Royal Statistical Society, LXXXIX (1926), 1-64.

